

**Electron Transfer in Proteins:
Theory and Experiment**

A Thesis
Submitted by
Ralf Langen

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Chemistry

Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, California

1995

Acknowledgements

I wish to express my sincere thanks to the following people:

Dr. Harry B. Gray, Dr. John H. Richards, and Dr. Arie Warshel, my advisors, for their guidance and support. Special thanks to Harry, who made everything possible by helping me to become a student at Caltech. To Jack, who provided me with a rich environment to perform my research. To Arie, who spent an extraordinary amount of time helping and teaching me, I learned a lot from you. I received the best training possible for a graduate student from these advisors.

Dr. Jay R. Winkler for his instructions and helpful discussions.

Cynthia Kiser, Neil Farrow, Claire Slutter, Jack Mizoguchi, and Phoebe Ray for providing a very helpful and exciting environment. Juris Germanas for training me, Danny Casimiro for many helpful discussions throughout my time at Caltech.

I-Jy Chang and Torbjorn Pascher for assistance with the laser experiments.

All members of the Gray group, especially Gary Mines, Tim Karpishin, Mark McClesky, Don Low, Deborah Wuttke, Mark Grinstaff, Kara Bren, and Jorge Colon for many helpful discussions.

All members of the Warshel group, especially Thomas Schweins, Z.T. Chu, Greg King, Luigi Manna, Mike Fothergill, and Fred Lee for sharing the lab, technical advice and laughter.

My parents and my wife for their support

ABSTRACT

This thesis consists of two main parts and one appendix. The first part (Chapters 2 and 3) contains mainly experimental work on electron transfer in ruthenium-modified azurin mutants. The second part describes theoretical studies on the control of reduction potentials in cytochromes and ferredoxins. The appendix presents calculations on the reaction mechanism for the GTP-hydrolysis in Ras p21.

In the experimental part electron tunneling through the β -sheet in *pseudomonas aeruginosa* azurin is discussed. In agreement with theoretical consideration it is found that electron tunneling through a β -strand is quite efficient. It is argued that the efficiency is due to the highly covalent nature of the bridge. Furthermore, the effects of the weakly coordinated M121 ligand and the strongly coordinated C112 ligand on the coupling are compared in Chapters 2 and 3. It is found that indeed the C112 ligand facilitates electron transfer, although the rate enhancement in the experiments does not appear to be much more than an order of magnitude.

The theoretical section (chapters 4-6) presents microscopical calculations of electrostatic effects found to be important for the control of reduction potentials. Chapter 4 analyzes the effect of the N52I mutation in cytochrome *c*. The calculations suggest that the change in reduction potential is due to an electrostatic effect and dominated by the loss of the N52 dipole, which compensates for the displacement of an internally bound water molecule. In chapter 5 it is concluded that the change in reduction potential for the M80H axial ligand replacement in cytochrome *c* is due to the effect of the ligand bonding rather than an electrostatic effect. The enormous variations in the reduction potentials of iron-sulfur clusters in several proteins are being investigated in chapter 6. The potentials mainly appear to be controlled by the number of hydrogen bonds. However, in some cases the effect of nearby water molecules need to be considered.

The calculations presented in the appendix show that a previously proposed mechanism in Ras P21 does not appear energetically feasible (reprint 1). In reprint 2 an alternative mechanism is proposed.

Table of Contents

Acknowledgements	ii
Abstract	iv
Table of Contents	v
List of Figures	vii
Chapter 1. Electron Transfer in Biology.	1
Chapter 2. Electronic Coupling through the M121 Strand in Azurin.	38
Chapter 3. Electronic Coupling through the C112 Strand in Azurin.	83
Chapter 4. Effect of the Asn52Ile Mutation on the Redox Potential of Cytochrome <i>c</i> .	119
Chapter 5. Simulation of the Change in Reduction Potential for the M80H Mutation in Cytochrome <i>c</i> .	157
Chapter 6. Protein Control of Iron Sulfur Cluster Redox Potentials.	174
<u>Appendix</u> - Calculations on the GTP Hydrolysis Mechanism in Oncoprotein p21.	188

- Reprint 1: Langen, R., Schweins, T, and Warshel, A. On the Mechanism of Guanosine Triphosphate Hydrolysis in Ras p21 Proteins. (1992) *Biochemistry* , **31**, 8691-8696 192
- Reprint 2: Schweins, T., Langen, R., and Warshel, A. Why have mutagenesis experiments not located the general base of Ras p21? (1994) *Nature Structural Biology*, **1**, 476-484 199

LIST OF FIGURES

Chapter 1

Figure 1.1. The Gibbs free-energy parabolas for the donor and acceptor state given as a function of a one-dimensional reaction coordinate.	23
1.2. Definition of the reorganization energy and the activation free energy using the Gibbs free-energy parabolas.	25
1.3. Logarithm of rate constant as a function of the driving force.	27
1.4. Decay per bond (ϵ_i^B) is plotted vs. energy.	29
1.5. Inverted behavior for D-A systems that are bridged by a steroid spacer.	31
1.6. $\ln k$ as a function of $-\Delta G^0$ for ruthenated cytochrome <i>c</i> derivatives.	33
1.7. Positions of ruthenium attachment sites in previous experiments in cytochrome <i>c</i> .	35
1.8. Electron transfer rate constants as function of distance and tunneling length for cytochrome <i>c</i> mutants.	37

Chapter 2

Figure 2.1. Plots of tunneling path lengths vs. β -carbon separation for idealized peptide secondary structures.	55
2.2. Backbone structure of azurin illustrating the location of the three surface His residues on the β -strand extending from the M121 ligand of the Cu atom.	57
2.3. Schematic drawing of the flash-quench and photoinduced techniques.	59
2.4. Representative kinetics for the H122 mutant in azurin.	61
2.5. Representative kinetics for the H124 mutant in azurin.	65

2.6. Representative kinetics for the H126 mutant in azurin.	69
2.7. Concentration dependence of the electron transfer kinetics for the H126 mutant in azurin.	74
2.8. Temperature dependence of the electron transfer kinetics for the H124 mutant in azurin.	76
2.9. Sequence of the ruthenium-containing fragment for the tryptic digest of the H124 in azurin.	78
2.10. Plot of $\log k_{\max}$ vs. metal-metal separation distance in Ru-modified azurin.	80
2.11. Comparison of tunneling length, distances, and rates in azurin and cytochrome <i>c</i> .	82

Chapter 3

Figure 3.1. Backbone structure of azurin showing the surface His residues modified in Chapter 2 and Chapter 3.	98
3.2. Representative kinetics for the H109 mutant in azurin.	100
3.3. Representative kinetics for the H107 mutant in azurin.	105
3.4. Concentration dependence of electron transfer rate constants for the H107 mutant in azurin.	110
3.5. The logarithm of the rate constants from electron transfer reactions in azurin mutants plotted versus metal-metal distance.	112
3.6. The logarithm of the rate constants from electron transfer reactions is plotted versus tunneling length.	114
3.7. Plot of tunneling length as a function of metal-metal separation in ruthenium-modified azurin and cytochrome <i>c</i> .	116
3.8. Plot of $\log(k_{\max})$ vs. metal-metal separation for Ru-modified azurin and cytochrome <i>c</i> .	118

Chapter 4

Figure 4.1. Thermodynamic cycle that relates the redox potentials of the native and the mutant cytochrome <i>c</i> .	146
4.2. Cyclic voltammogram of Asn52Ile cytochrome <i>c</i> .	148
4.3. Ribbon drawing of the polypeptide backbone chain of yeast iso-1-cytochrome <i>c</i> .	150
4.4. Drawings of the immediate vicinity of a) Asn-52 in wild-type yeast iso-1 cytochrome <i>c</i> and of b) Ile-52 in the Asn52Ile mutant protein.	152
4.5. Convergence of the average electrostatic energy (ΔG_{elec}) as function of simulation time for the Asn52Ile mutation in cytochrome <i>c</i> .	154
4.6. The relationship between the energetics of the reduced and oxidized cytochromes and the coordinate of the protein and the solvent.	156

Chapter 5

Figure 5.1. Redox potential of M80H cytochrome <i>c</i> in comparison to wild-type cytochrome <i>c</i> and cytochrome <i>b5</i> .	169
5.2. Convergence of the average electrostatic energy (ΔG_{elec}) as function of simulation time for the M80H mutation in cytochrome <i>c</i> .	171
5.3. Comparison of the solvation energy contributions for the oxidation of cytochrome <i>c</i> and cytochrome <i>b5</i> .	173

Chapter 6

Figure 6.1. Contributions to the differential solvation energies of the $[Fe_4S_4(S-cys)_4]^{-2/3}$ clusters of <u>Bt</u> Fd, <u>Pa</u> Fd, <u>Av</u> FdI, and <u>Cv</u> HiPIP.	183
6.2. Calculated and experimental redox potentials of <u>Bt</u> Fd, <u>Pa</u> Fd, <u>Av</u> FdI, and <u>Cv</u> HiPIP.	185
6.3. The protein/Langevin dipole system for <u>Pa</u> Fd and <u>Av</u> FdI.	187

Chapter 1

Electron Transfer in Biology

BACKGROUND

Electron transfer (ET) plays a key role in a number of biological processes, such as photosynthesis, oxidative phosphorylation, nitrogen fixation, and metabolism [1-6]. This explains the significant attention the field has received over the past 30 years from theoreticians and experimentalists. Especially in the last decade our knowledge of biological ET has increased significantly. The determination of the crystal structure of the reaction center protein in photosynthetic bacteria [7], together with ultrafast laser spectroscopy has laid the foundation necessary to obtain a detailed mechanistic picture of the processes involved. Significant insight has been gained into how nature has optimized the efficacy of the system. Equally important for our understanding of ET were advances made in the field of molecular biology. With the advent of site-directed mutagenesis it is now possible to generate specific changes in the protein matrix in order to study the effects single amino acids have on reduction potentials and/or ET kinetics. One of the central issues is the control of the ET rates by nuclear and electronic factors, which will be discussed in detail in the latter section of this chapter. In particular, the issue of pathway dependence of ET has received great attention. Not only does site-directed mutagenesis provide the option to modify amino acids in the putative ET path, it also creates the opportunity to introduce specific ligands at various positions that can be modified with artificial redox-sites. The latter strategy has been used extensively in the work described in this thesis.

BASIC ELECTRON TRANSFER THEORY

An electron transfer reaction is best described as the transition of an electron from a donor (D) to an acceptor (A). Such a reaction can either be intramolecular or intermolecular. The ET theory for both processes is very similar. However in intermolecular reactions one also has to consider the diffusion of D and A to form of the D-A precursor complex.

If one describes ET using classical transition state theory, the first-order rate constant is given by

$$k_{ET} = \kappa_{el} \nu_n \exp\left[\frac{-\Delta G^\ddagger}{k_b T}\right] \quad (1)$$

where ν_n is the frequency of motion along the reaction coordinate ($\sim 10^{13}\text{s}^{-1}$) due to the nuclear motions, ΔG^\ddagger is the Gibbs energy of activation, k_b is the Boltzmann constant, T is temperature. κ_{el} is the transmission coefficient, which is a measure of the probability of an ET reaction to occur, once the transition state is reached. It is this factor that lets us distinguish two types of ET. For a κ_{el} close to unity, ET occurs almost every time the nuclear configurations are at or near the transition state geometry. In this case ET is called adiabatic. The other extreme ($\kappa_{el} \ll 1$) is often found in proteins where ET between D and A occurs over distances often exceeding 10 Å. Those reactions are said to be nonadiabatic.

Typical ET reactions in solution depend on an enormous number of nuclear coordinates, which greatly complicates the analysis of such reactions. In the pioneering work by Rudolph Marcus [8-12] this problem was circumvented by the definition of a one-dimensional reaction coordinate. It can be shown that, if one uses Gibbs free energy space, the energy profiles for a donor and an acceptor state can be well approximated by parabolas. An example is given in Fig. 1.1. In Marcus theory, the curvature of the

parabolas is taken to be the same. In order to satisfy the Franck-Condon principle and the conservation of energy, the thermal transition from the reactant to the product surface can only occur at or around C^* . The representation of ET by two parabolas lead to an important relation between the activation free energy and the reorganization energy (λ). The reorganization energy can be defined as the change in free energy required to move on the reactant surface from its equilibrium nuclear coordinate value to the equilibrium nuclear coordinate value of the product surface without actual ET (Fig. 1.2). In the case of a reaction with a $\Delta G^0=0$ (for example a self-exchange reaction) it follows from the behavior of parabolas that

$$\Delta G^* = \frac{\lambda}{4} \quad (2)$$

In reactions for which ΔG^0 is not equal to 0 one usually simply moves the position of the product surface vertically (Fig. 1.2). Again, it follows from the geometry of intersecting parabolas that

$$\Delta G^* = \frac{(\lambda + \Delta G^0)^2}{4\lambda} \quad (3)$$

Equations (3) and (1) are combined to give the classical Marcus equation

$$k_{ET} = \kappa_{el} \nu_n \exp\left[\frac{-(\lambda + \Delta G^0)^2}{4\lambda k_b T}\right] \quad (4)$$

It can be seen from equation (4) that the rate of electron transfer increases with increasing driving force. It reaches a maximum, when $-\Delta G^0 = \lambda$. The surprising prediction from this equation is that as the driving force increases even further, the rate is expected to drop. This high-driving force regime is termed the inverted region (Fig. 1.3).

* During the ET the much heavier and slower nuclei have no time to change either their position or their momentum.

The reorganization energy is usually the sum of two components, an inner sphere part (λ_{in}) and a solvent contribution (λ_{out}). Often λ_{in} is relatively small compared to the solvent reorganization, and the overall reorganization energy is dominated by λ_{out} . Several expressions for the solvent reorganization energy have been derived using a dielectric continuum model. Those calculations show that the reorganization energy is larger for polar than for non polar solvents as the latter ones do not reorient as much in response to changes in the electric fields. The calculations also predict that the reorganization energy is a function of D-A separation. While the use of a continuum model to describe bulk solvent properties seems appropriate, a reliable estimate of the reorganization energies in biological molecules should include the nature of the peptide matrix into the calculations. Such calculations have been done for cytochrome *c* using a microscopic model [13]. More discussion of these microscopic approaches and more recent related microscopic simulations of ET in proteins and solutions are reviewed in reference [1].

So far the discussion of ET has neither considered the electronic coupling between donor and acceptor, nor has it accounted for the vibrational states accessible to donor and acceptor. This can be included by using the Fermi golden rule expression

$$k = \frac{2\pi}{\hbar} H_{DA}^2 (FC) \quad (5)$$

where H_{DA} is the electronic coupling matrix element of the reactants' electronic state with the products'. (FC) could be described as a Franck-Condon weighted density of states. It represents the sum of individual products of donor and acceptor vibrational wave functions, weighted by the Boltzmann probability. In a fully quantum mechanical analysis this sum would need to be taken over all vibrational modes. However, since the solvent vibrations usually occur at low frequencies they can be treated classically. Using such an approximation it is possible to obtain Eq. (6).

$$k_{ET} = \frac{2\pi}{\hbar} H_{DA}^2 \sum_{v_r, v_p} \frac{1}{\sqrt{4\pi\lambda_{out}k_bT}} \exp\left[-\frac{(\lambda_{out} + \Delta G^0 + \Delta\epsilon_v)^2}{4\lambda_{out}k_bT}\right] (FC)_v P(v_r) \quad (6)$$

where $(FC)_v P(v_r)$ is the Boltzmann-weighted sum of solvational Franck-Condon factors.

In the high temperature limit equation (6) reduces to

$$k_{ET} = \frac{2\pi}{\hbar} H_{DA}^2 \frac{1}{\sqrt{4\pi\lambda k_bT}} \exp\left[-\frac{(\lambda + \Delta G^0)^2}{4\lambda k_bT}\right] \quad (7)$$

This expression (eq. 7) still retains the general feature that the rate of reaction is maximal, when $-\Delta G^0 = \lambda$.

ELECTRON TUNNELING PATHWAYS

ET in biological molecules often involves charge transport over long distances [1-6]. In these cases ET cannot be considered adiabatic. The reaction only occurs very rarely, when the transition state geometry is reached, and equations 5-7 should be used. In all of those equations, the rate of ET depends on the square of the electronic coupling matrix element, H_{DA} . In a simple two-state system only containing the donor and acceptor orbitals, H_{DA} would be a function of the overlap of those orbitals. At distances found in biological processes this overlap would be vanishingly small, and therefore the rate constant would have to be extremely low. Thus, it appears to be clear that the intervening medium between donor and acceptor (bridge) has an important contribution in mediating ET. Typically, the energies of the bridge orbitals are quite different from the donor and the acceptor orbitals, and the electron transfer appears to occur by tunneling through the protein matrix.

Several models for electron tunneling in biological molecules have been proposed. The simplest is the square barrier model, in which the protein is viewed as a fully homogeneous bridge, and H_{DA} is simply a function of distance [11, 16, 17].

$$H_{DA} = H_{DA}^0 \exp[-\beta(d-d_0)/2] \quad (8)$$

where d is the D-A separation. This simple model often provides a surprisingly good description for biological ET (mainly for the reaction center [17]). However, the assumption that the bridge is homogeneous cannot always be made. Several experiments have shown that the electronic coupling depends on the make-up of the bridge [18-20]. An example will be presented in this thesis in chapters 2 and 3.

In many theoretical studies, the inhomogeneity of the bridge has been addressed. Siddarth and Marcus [21-25] have used an artificial-intelligence search to evaluate the relevant bridge amino-acids, and applied extended Hückel calculations to evaluate the electronic coupling elements. Gruschus and Kuki [26] have developed an inhomogeneous aperiodic lattice theory. The model, that will be discussed in more detail in this chapter is the Beratan&Onuchic (B&O) pathway model [26-32], since it was used to a great extent in the design and the interpretation of the experiments presented in this thesis.

Beratan and Onuchic developed a model that considers physical tunneling pathways. These pathways are best defined as a collection of interacting bonds extending the donor into the acceptor wave functions. Such a path can consist of a combination of covalent bonds (B), hydrogen bonds (H), and through-space-jumps (S) between atoms. For one physical pathway the tunneling matrix element can be written as

$$t_{DA} = \text{prefactor} \prod \epsilon_i^B \prod \epsilon_j^S \prod \epsilon_k^H \quad (9a)$$

with

$$\text{prefactor} = \frac{\beta_A \beta_B \beta_1}{(E - \alpha_L^1)(E - \alpha_R^1) - b_1^2} \quad (9b)$$

The prefactor in equation (9) contains the information about the beginning and the end of the pathway with β_D (β_A) being the coupling between the donor (acceptor) and the first (last) bond of the bridge. The remaining terms in the prefactor describe the propagation of the donor wave function into the first bond of the bridge. The electronic coupling between the atomic orbitals of bond 1 is β_1 . α_R^1 and α_L^1 are the respective energies of these orbitals and E is the energy of the tunneling electron. The ϵ 's in equation 9 are the individual decay factors for the propagation of the donor wave function through bonds (B), through hydrogen bonds (H), and through space (R).

Through-bond tunneling can occur via two different mechanisms. It can be mediated by the bonding orbitals of the bridge, in which case the electron transfer is referred to as hole tunneling (h), conversely tunneling can occur through antibonding orbitals of the bridge, which is called electron tunneling (e). Which one of these mechanisms is dominant depends on which bridge levels are energetically closest to the energy of the tunneling electron

$$\epsilon^B = \epsilon^e + \epsilon^h \quad (10)$$

with

$$\epsilon^h = \frac{\gamma}{E - \alpha_{bond}} \quad (11)$$

$$\epsilon^e = -\frac{\gamma}{E - \alpha_{anti}} \quad (12)$$

A plot of ϵ^h and ϵ^e is shown in Fig. 1.4. Due to the energetics in proteins, the electron transport seems to be mainly dominated by hole tunneling. The pathway model uses a constant of 0.6 per bond as decay factor.

It is important to note that the bond types considered in the pathway model are exclusively σ -bonds. Although π -interactions could lead to a rate enhancement, it appears to be a good approximation to focus on the σ -bonds in proteins, which usually do not

contain extended π -systems. It appears that the possible rate enhancement due to the good coupling in a π -system is compensated by the unfavorable σ - π mixing, that has to occur when the electron enters the π -system from the σ -pathway. Similarly, the exit from the π -system back onto the σ -pathway is also unfavorable. This interpretation is in good agreement with experimental data [33].

Decay through hydrogen bonds is treated similar to the through σ -bond decay. In the pathway model a hydrogen bond can be considered as an extended σ -bond. The parameterization is such that a hydrogen bond at equilibrium distance is equivalent to two covalent bonds corrected for the displacement from the equilibrium distance R_H^{eq} .

$$\epsilon^H = (\epsilon^B)^2 \exp[-\beta^H(R - R_H^{eq})] \quad (13)$$

where R is the actual hydrogen bond length, and β^H is a decay constant.

While the parameterization of the through-bond decay could make use of experimental data, this is very difficult for the through-space (non covalent) orbital interactions, therefore theoretical estimates have to be used. It is again reasonable to assume that the bond (or resonance) integrals are proportional to the overlap between the two bonding orbitals. If, as before, orientation or symmetry effects are neglected the overlap is expected to have an exponential decay with distance.

$$\epsilon^H = (\epsilon^B) \exp[-\beta^B(R - R_B^{eq})] \quad (14)$$

An important quantity that can be derived from the B&O-type calculations is the tunneling length (σ). Once the tunneling matrix element has been determined for a given pathway, one can ask how many purely covalent bonds would result in the same coupling. If one multiplies this number by 1.4\AA , one obtains the tunneling length. Examples will be given in a latter part of this chapter as well as in Chapters 2 and 3.

ELECTRON TRANSFER IN SMALL MOLECULES

From the above discussion it is expected that fundamental parameters controlling the rate of ET are distance, Franck-Condon factors, and the electronic coupling. Very helpful for our understanding of ET in proteins have been studies on small molecules. For example, one of the crucial predictions arising from the theoretical considerations described above was the existence of the inverted region. The experimental test of this prediction was not easy, due to problems arising from diffusion limitations. Nevertheless, it was possible to convincingly demonstrate inverted behavior using models with covalently linked D-A systems, in which a steroid bridge connects various acceptors group to a biphenyl donor [34, 35]. The results are shown in Fig. 1.5. Since then, the inverted effect has been observed in several other systems ([36-38] for example).

As far as the distance dependence of the ET rate in small molecules is concerned, the per-distance decay model and the per-bond decay model typically agree equally well with the data. This may be attributed to the fact that in most experiments the distance and the number of bonds cannot be varied independently. Often, the number of intervening bonds is proportional to the separation distance. Nevertheless, in some experiments those quantities were varied independently. Maybe one of the best hints that a through-bond tunneling model might describe the physics of ET reactions best came from comparison of rates of isomeric decalin-bridged D-A complexes. The two isomers have quite different center-center separation distances (difference=6.3 Å), yet the observed ET rate constants are almost identical [35].

ELECTRON TRANSFER IN METALLOPROTEINS

In view of the above discussion an important question arises for biological ET. How does the nature of the bridge influence the electronic coupling between donor and acceptor?

Possibly the most powerful methodology to address the issue of specific tunneling pathways represents the study of intramolecular ET, in which the positions of D and A are precisely defined. Numerous biological systems satisfy these criteria. For example, it is possible to study ET in the reaction center which has "built-in" D and A sites. In some cases, when a protein contains only one "natural redox center," it has been possible to take advantage of a disulfide bond present in the protein. For example, in azurin the disulfide bond can be reduced with the help of pulse radiolysis [39-42]. The reduced disulfide now represents a good reductant that can reduce the active site copper. Using site-directed mutagenesis techniques it is possible to test a given tunneling pathway by varying the bridge amino acids. A drawback of this methodology is the difficulty to vary the D-A separation. This problem can be circumvented by the covalent attachment of an artificial redox-site onto a given protein.

Ruthenium complexes are commonly used for this purpose (for a review see [43]). The first work on ET reactions in ruthenium-modified proteins was performed in horse heart cytochrome *c* [44, 45]. The synthesis involves the coordination of a pentaammineruthenium complex to the native H33. Using the same methodology, several other native or engineered surface histidine residues have been modified since [43]. In those experiments photochemically induced ET occurs at relatively low driving force. This represents a problem, since one is often interested in the maximum rate constant (see equation 4, 6, or 7). For an extrapolation of the maximum rate constant, however, it is necessary to know the reorganization energy (λ). In principal, one could obtain this information from temperature and driving-force studies. However, the evaluation of the

reorganization energy from temperature dependence is complicated by the temperature dependence of the driving force itself. Furthermore, it is difficult to obtain a good estimate from driving force dependence studies in the low driving force regime. Since changes in the rate constant with varying driving force are quite independent of λ in this regime, only little information about λ can be extracted from those experiments. By substituting the native Fe in cytochrome *c* with Zn, however, it is possible to move the ET reactions into the high-driving-force regime (see for example [46, 47]). Results from several such studies in cytochrome *c* are shown in Fig. 1.6. In all cases the reorganization energy appears to be approximately 1.2 eV. Since all of the derivatives are at different positions the data suggest that the distance dependence of the reorganization energy mentioned earlier must be very small.

Another method for moving the ET reaction into the high-driving-force regime is to use Ru(bpy)₂(im)-His derivatives [48]. For proteins like azurin and cytochrome *c* such derivatizations allow the measurement of ET rates for which $-\Delta G \approx \lambda$. Using equation 7 it is possible to extract the electronic coupling matrix element directly from the observed rate constants. This approach has been used to test the predictions obtained from the B&O pathway model described above. For cytochrome *c*, the observed and predicted rates are in very good agreement, while a square-barrier tunneling model does not satisfactorily describe the measured rates [18-20]. This is especially true for the H72 derivative [19, 20], whose tunneling pathway contains a long through-space jump (Fig. 1.7). As predicted, this mutant exhibits a much slower ET rate constant than expected from a square-barrier tunneling model with a β of 1.4/Å (Fig. 1.8). When the data are plotted versus tunneling length, the through-space contact is accounted for by a very long tunneling length for the through space contact (10.6 bonds \times 1.4 Å/bond = 14.8 Å).

It appears that the success of the pathway model in predicting the electronic coupling in cytochrome *c* is partly due to the fact that dominant electron tunneling pathways can be

identified. This is also the case for ET rates measured in azurin (Chapters 2 and 3 of this thesis), which can very well be interpreted in terms of the B&O pathway model. In myoglobin, however, the pathway model identifies many equivalent pathways, and this complicates the analysis, because interferences need to be considered [49]. Currently a model is being developed that takes the possible interferences from the individual pathways into account. It should be pointed out here that the approach introduced by Siddarth and Marcus considers such interferences and gives agreement with experiment [25].

ELECTROSTATIC CONTRIBUTIONS TO REDUCTION POTENTIALS AND REORGANIZATION ENERGIES IN ET PROTEINS

The discussion so far has shown that electronic coupling and the Franck-Condon Factors are key factors that determine the rate of ET. In this section it will be discussed how one can evaluate the electrostatic effects contributing to the Franck-Condon factors. From equations 4-7, it becomes apparent that one of the requirements is the knowledge of reorganization energy and free energy. The free-energy change can be expressed in terms of the reduction potentials (E) of D and A.

$$\Delta G = zF(E_D - E_A) \quad (16)$$

where z is the number of electrons transferred, and F is the Faraday constant. What are the factors that control the reduction potential and the reorganization energy of a protein? One of the most important contributions arises from the electrostatic environment around a given cofactor. In order to evaluate this effect one needs to determine the self-energy, which is the free energy difference associated with moving a charged group from vacuum to a given medium with dielectric constant ϵ . In principal, such a process can be calculated using the Born formula [50]

$$\Delta G = -322 \frac{Q^2}{2a} \left(1 - \frac{1}{\epsilon}\right) \quad (17)$$

where Q is the charge of the group, and a is its effective radius. Unfortunately, ϵ is a macroscopic constant and it is difficult to define ϵ in an inhomogeneous medium such as a protein. This leaves only two options. One either has to use a microscopic model that does not require the knowledge of ϵ , or one has to find an estimate for ϵ , that is valid for a given region of the protein. In the case of a protein one needs to consider at least two different regions, the polypeptide and the bulk solvent. This feature is considered in the Tanford-Kirkwood (TK) model [51], in which the protein is assigned a low dielectric constant, surrounded by a solvent of high dielectric constant. If one modifies the original TK model by adding the missing self-energy terms one can write (for a detailed discussion see [52, 57]):

$$\Delta G = -166Q^2 \left[\frac{1}{a} \left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{1}{b} \left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_{out}}\right) \right] \quad (18)$$

where b is the radius of a sphere with dielectric constant ϵ_{in} (small). ϵ_{out} is the dielectric constant of the surrounding solvent (large). Although this model gives quite reasonable results for the treatment of charged surface residues, it has serious limitations when internal charges are considered. Ionized groups are frequently observed in the interior of proteins, and it seems that, at least in most cases, they can be quite stabilizing. According to the TK model, this would not be possible, because the energy of bringing a charge from the high-dielectric medium to a low-dielectric medium would be substantial. Somewhat surprising might be that even if one assumes a dielectric constant for the protein much larger than $\epsilon_{out} = 80$ (usually assumed for water), according to equation 18 there will be little or no effective stabilization of a charge in a protein (note the reciprocal values of ϵ in equation 18). Although one could in principal adjust the value of a to an unrealistically small number, it appears that a physically meaningful description needs to

consider the protein dipoles explicitly. Such a microscopic approach has been implemented into several macroscopic models now (see for example [53, 54]).

Another limitation of the model is the assumption that the protein can be considered as a sphere. Since the energy is a quite steep function of distance, it would be desirable to consider the actual geometry of a real protein[†]. An improvement in this regard was made by the introduction of a discretized continuum approach [55, 56]. This model superimposes a grid onto the protein and solvent system and assigns a dielectric constant for each grid point. With such an assumption it becomes possible, to numerically solve the Poisson equation.

$$-\nabla^2 \Phi(r) = \frac{\rho(r)}{\epsilon_r} \quad (19)$$

Here $\Phi(r)$ is the electrostatic potential and $\rho(r)$ is the free charge density. This method and different variants thereof are widely being using today (see for example [53, 54]).

A microscopic approach, that does not require any information on the clearly site-dependent dielectric constant has been developed by Warshel [57, 58]. This model is based on the assumption that as soon as all the relevant electrostatic contributions are included in the calculations, one does not need to know the value of ϵ at all. The important terms in the model are charge-charge interactions, charge-induced dipole interactions, and the charge-water interaction. The charge-water interactions inside the protein are evaluated using Langevin-type dipoles. The solvent outside the protein is modeled by a continuum approach. For a more detailed discussion of the method see reference [57, 58] or Chapter 4.

This approach has been used to calculate the reorganization energy (λ) in cytochrome c [13]. The calculations were based on the x-ray crystallographic structures of the reduced

[†] Note that the Tanford-Kirkwood model was developed before x-ray crystallographic structures became available.

and oxidized protein, and it has been shown that λ is very small consistent with its catalytic function. It has been of particular interest in those calculations, to test the validity of equations 2 and 3 using a microscopic model. As stated above, those equations have been derived using a continuum model that assumes a linear response of the solvent. This does not necessarily have to be the case in the protein. Nevertheless, the microscopic calculations are consistent with equations 2 and 3 suggesting that to a good approximation the average response of the protein dipoles can also be viewed as linear with the local field.

Similarly, the same model was used to evaluate the reduction potentials of proteins [59-61]. Crucial for those calculations is the appropriate choice of a reference system. For example, calculations on the reduction potential in cytochrome *c* compared the self-energies of a heme group in the protein to the corresponding energy in water. The calculations show that the high reduction potential in cytochrome *c* (when compared to the model compound in water) can be explained by the low polarity (i.e. lack of charge stabilizing interaction) of the protein. This point has already been recognized by the Kassner macroscopic approach [62]. Here, we have an example where a microscopic and a macroscopic approach that assumes a low dielectric constant for a protein lead to the same conclusion. This agreement should generally be expected for environments that destabilize a given charge. However, often the protein matrix stabilizes a charge more strongly than water, which cannot be understood in terms of the TK model. This is the case in some iron-sulfur cluster containing proteins as will be argued in Chapter 6.

BIBLIOGRAPHY

1. Warshel, A., and Parson, W.W., *Ann. Rev. Phys. Chem.* **42**, 279 (1991)
2. Gray, H.B., and Malmström, B.G., *Biochemistry* **28**, 7499 (1989)
3. Mayo, S.L., Ellis, W.R., Crutchley, R.J., and Gray, H.B., *Science* **233**, 948 (1986)
4. Adman, E.T., *Biochem. Biophys. Acta* **549**, 107 (1979)
5. Sigel, H., and Sigel, A.: *Metal Ions in Biological Systems*, New York: Dekker; 1991.
6. Bertini, I., Gray, H.B., Lippard, S.J., and Valentine, J.S.: *Bioinorganic Chemistry*, Mill Valley, CA: University Science Books; 1994.
7. Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H.J., *J. Mol. Biol.* **80**, 385 (1984)
8. Marcus, R.A., *J. Chem. Phys.* **24**, 966 (1956)
9. Marcus, R.A., *Annu. Rev. Phys. Chem.* **15**, 155 (1960)
10. Marcus, R.A., *J. Chem. Phys.* **43**, 679 (1965)
11. Marcus, R.A., and Sutin, N., *Biochim Biophys. Acta* **811**, 265 (1985)
12. Marcus, R.A., and Sutin, N., *Comments Inorg. Chem.* **5**, 119 (1986)
13. Churg, A.K., Weiss, R.M., Warshel, A., and Takano, T., *J. Phys. Chem.* **87**, 1683 (1983)
14. Marcus, R.A., *Faraday Disc. Chem. Soc.* **55**, 34 (1973)
15. Marcus, R.A., *J. Chem. Phys.* **81**, 4494 (1984)

16. Hopfield, J.J., *Proc. Natl. Acad. Sci* **71**, 3640 (1974)
17. Moser, C.C., Keske, J.M., Warncke, K., Farid, R.S., and Dutton, P.L., *Nature* **355**, 796 (1992)
18. Karpishin, T.B., Grinstaff, M.W., Komar-Panicucci, S., McLendon, G., and Gray, H.B., *Structure* **2**, 415 (1994)
19. Wuttke, D.S., Bjerrum, M.J., Winkler, J., and Gray, H.B., *Science* **256**, 1007 (1992)
20. Wuttke, D.S., Bjerrum, M.J., Chang, I-Jy, Winkler, J.R., and Gray, H.B., *Biochim. Biophys. Acta* **1101**, 168 (1992)
21. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **94**, 8430 (1990)
22. Siddarth, P., and Marcus, R.A., *J. Chem. Phys.* **94**, 2985 (1990)
23. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **96**, 3213 (1992)
24. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **97**, 2400 (1993)
25. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **97**, 13078 (1993)
26. Gruschus, J.M., and Kuki, A., *J. Phys. Chem.* **1993**, 5581 (1993)
27. Beratan, D.N., Betts, J.N., and Onuchic, J.N., *Science* **252**, 1285 (1991)
28. Beratan, D.N., Betts, J.N., and Onuchic, J.N., *J. Phys. Chem.* **96**, 2852 (1992)
29. Onuchic, J.N., and Beratan, D.N., *J. Chem. Phys.* **92**, 722 (1990)
30. Onuchic, J.N., Andrade, P.C.P., and Beratan, D.N., *J. Chem. Phys.* **95**, 1131 (1991)

31. Onuchic, J.N., Beratan, D.N., Winkler, J.R., and Gray, H.B., *Annu. Rev. Biophys. Biomol. Struct.* **21**, 349 (1992)
32. Skourtis, S.S., Regan, J.J., and Onuchic, J.N., *J. Phys. Chem.* **98**, 3379 (1994)
33. Casimiro, D.R., Richards, J.R., Winkler, J.R., and Gray, H.B., *J. Phys. Chem* **97**, 13073 (1993)
34. Closs, G.L., Calcaterra, L.T., Green, N.J., Penfield, K.W, and Miller, J.R., *J. Phys. Chem.* **90**, 3673 (1986)
35. Closs, G.L., Miller, J.R., *Science* **240**, 440 (1988)
36. Gould, I.R., Ege, D., Mattes, S.L., and Farid, S., *J. Am. Chem. Soc.* **109**, 3794 (1987)
37. Fox, L.S., Kozik, L.S., Winkler, J.R., and Gray, H.B., *Science* **247**, 1069 (1990)
38. McClesky, T.M., Winkler, J.R., and Gray, H.B., *J. Am. Chem. Soc.* **114**, 6935 (1992)
39. Faraggi, M., and Klapper, M.H., *J. Am. Chem. Soc.* **110**, 5753 (1988)
40. Klapper, M.H., and Faraggi, M.Q., *Quart. Rev. Biophys.* **12**, 465 (1979)
41. Farver, O., and Pecht, I., *J. Am. Chem. Soc.* **114**, 5764 (1992)
42. Farver, O., Skov, L.K., Pascher, T., Karlsson, G., Nordling, M., Lundberg, L.G., Vänngård, T., and Pecht, I., *Biochemistry* **32**, 7317 (1993)
43. Winkler, J.R., and Gray, H.B., *Chem Rev.* **92**, 369 (1992)
44. Nocera, D.G., Winkler, J.R., Yocom, K.M., Bordignon, E., and Gray, H., B., *J. Am. Chem. Soc.* **196**, 5145 (1984)

45. Winkler, J.R., Nocera, D.G., Yocom, K.M., Bordignon, E., and Gray, H.B., *J. Am. Chem. Soc.* **104**, 5798 (1982)
46. Elias, H., Chou, M.H., and Winkler, J.R., *J. Am. Chem. Soc.* **110**, 599 (1988)
47. Meade, T.J., Gray, H.B., and Winkler, J.R., *J. Am. Chem. Soc.* **111**, 4353-4356 (1989)
48. Chang, I.-J., Gray, H.B., and Winkler, J.R., *J. Am. Chem. Soc.* **113**, 7056 (1991)
49. Casimiro, D.R., Wong, L.-L., Colon, J.L., Zewert, T.E., Richards, J.H., Chang, I.-J., Winkler, J.R., and Gray, H.B., *J. Am. Chem. Soc.* **115**, 1485 (1993)
50. Born, M., *Z. Phys.* **1**, 45 (1920)
51. Tanford, C., and Kirkwood, J.G., *J. Am. Chem. Soc.* **79**, 5333 (1957)
52. Warshel, A., and Aqvist, J., *Ann. Rev. Biophys. Biophys. Chem.* **20**, 267 (1991)
53. Nicholls, A., and Honig, B., *J. Comp. Chem.* **12**, 435 (1991)
54. Gunner, M.R., and Honig, B., *Proc. Natl. Acad. Sci.* **88**, 9151 (1991)
55. Orthung, W.H., *Ann. NY Acad. Sci.* **303**, 22 (1977)
56. Warwicker, J., and Watson, H.C., *J. Mol. Biol.* **157**, 671 (1982)
57. Russel, S., and Warshel, A., *J. Mol. Biol.* **185**, 389 (1985)
58. Warshel, A., and Russel, S., *Quart. Rev. Biophys.* **17**, 283 (1984)
59. Churg, A.K., and Warshel, A., *Biochemistry* **25**, 1675 (1986)
60. Langen, R., Jensen, G.M., Jacob, U., Stevens, P.J., and Warshel, A., *J. Biol. Chem.* **267**, 25625 (1992)

61. Langen, R., Brayer, G.D., Berghuis, A.M., McLendon, G., Sherman, F., and Warshel, A., *J. Mol. Biol.* **224**, 589 (1992)
62. Kassner, R.J., *Proc. Natl. Acad. Sci* **69**, 2263 (1972)

Figure 1.1. The Gibbs free energy parabolas for the donor and acceptor state are given as function of the one-dimensional reaction coordinate. A is the equilibrium nuclear coordinate for the reactant (R) surface. B is the equilibrium nuclear coordinate for the product (P) surface. Thermally the transition from one surface to another can only occur close to point C.

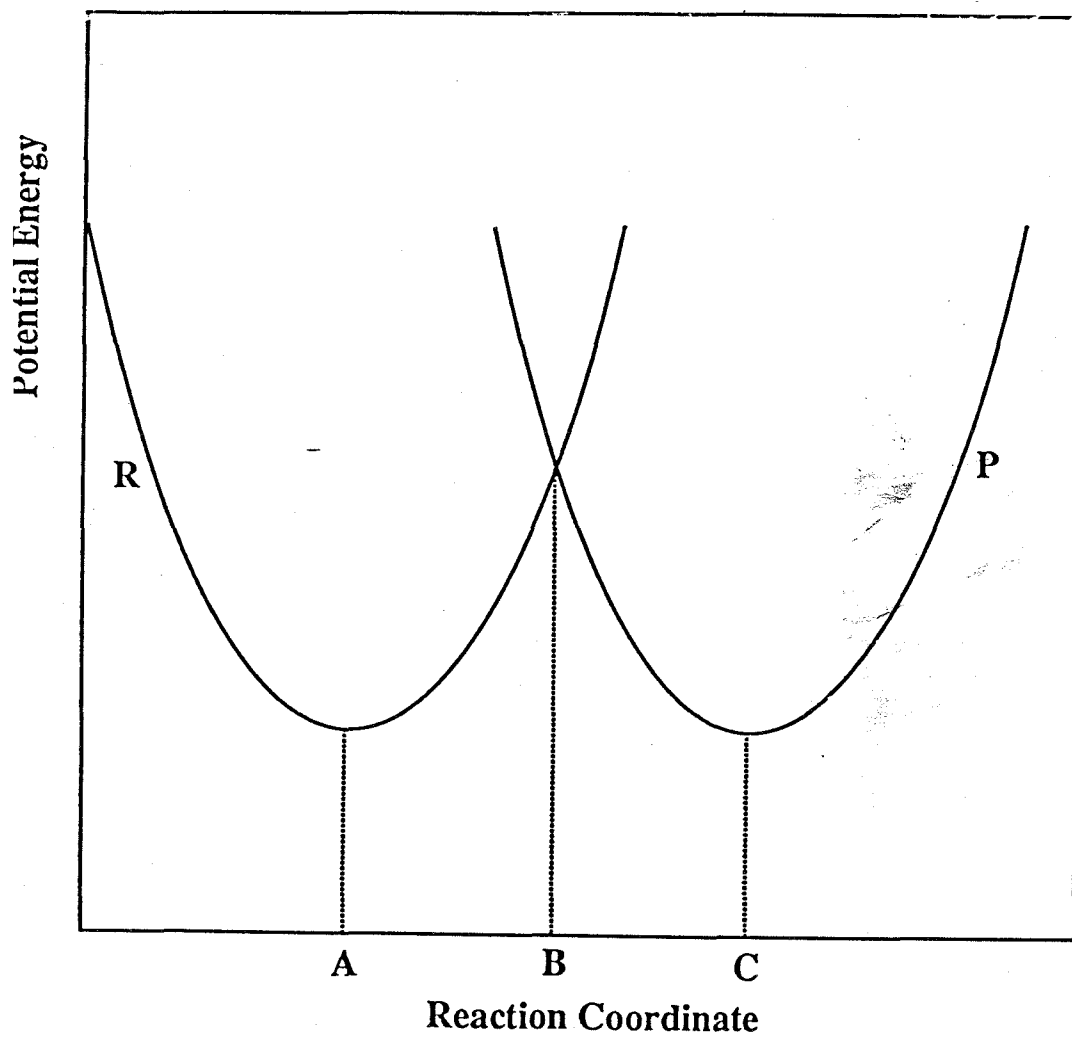


Figure 1.2. Diagrams showing the intersecting Gibbs parabolas for (A) a reaction with $\Delta G^0=0$, (B) $\Delta G^0<0$. The reorganization energy (λ), and the activation free energy (ΔG^*).

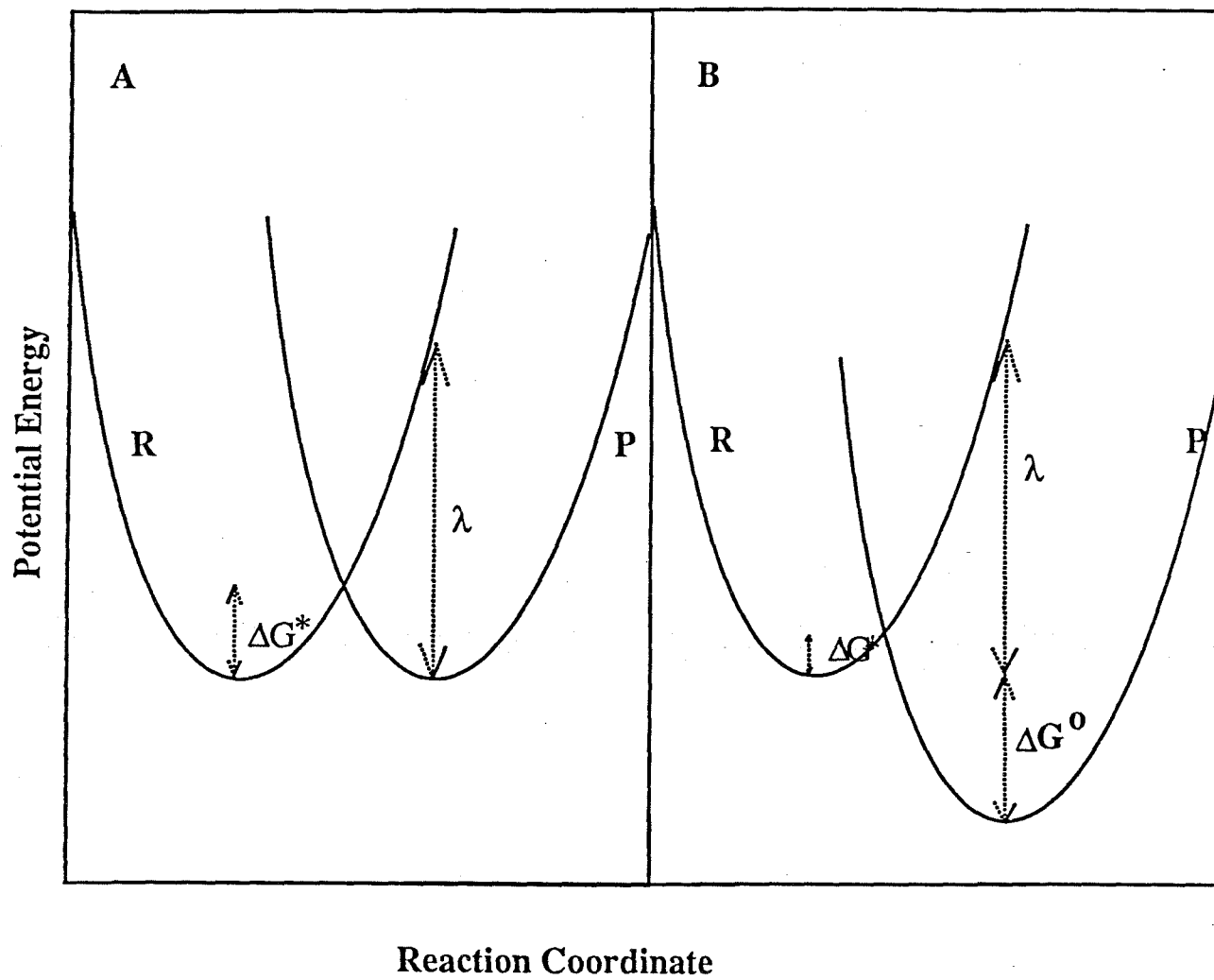


Figure 1.3. The dependence of the log of the rate constant as function of the driving force is shown schematically. The maximum rate is at $-\Delta G^0 = \lambda$.

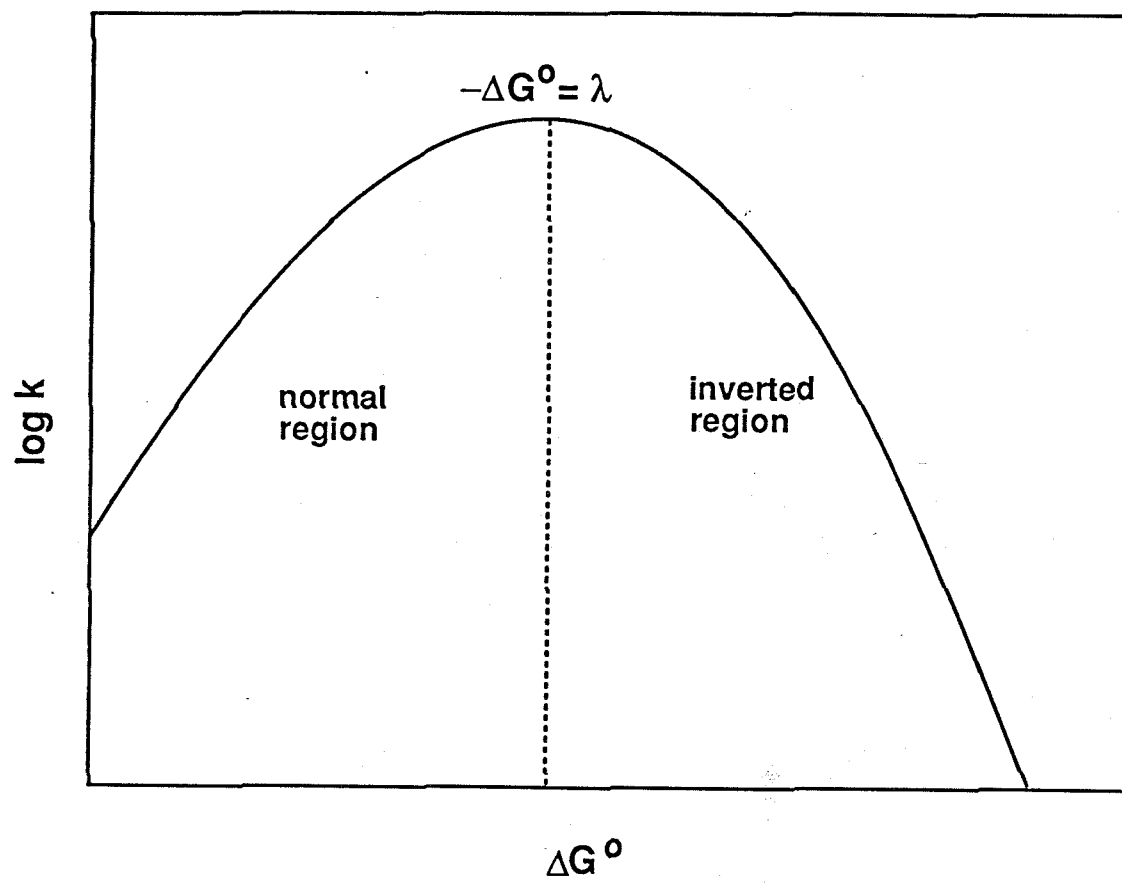


Figure 1.4. Decay per bond (ξ_1^B) is plotted vs. E. Hole and electron transfer limits are shown.

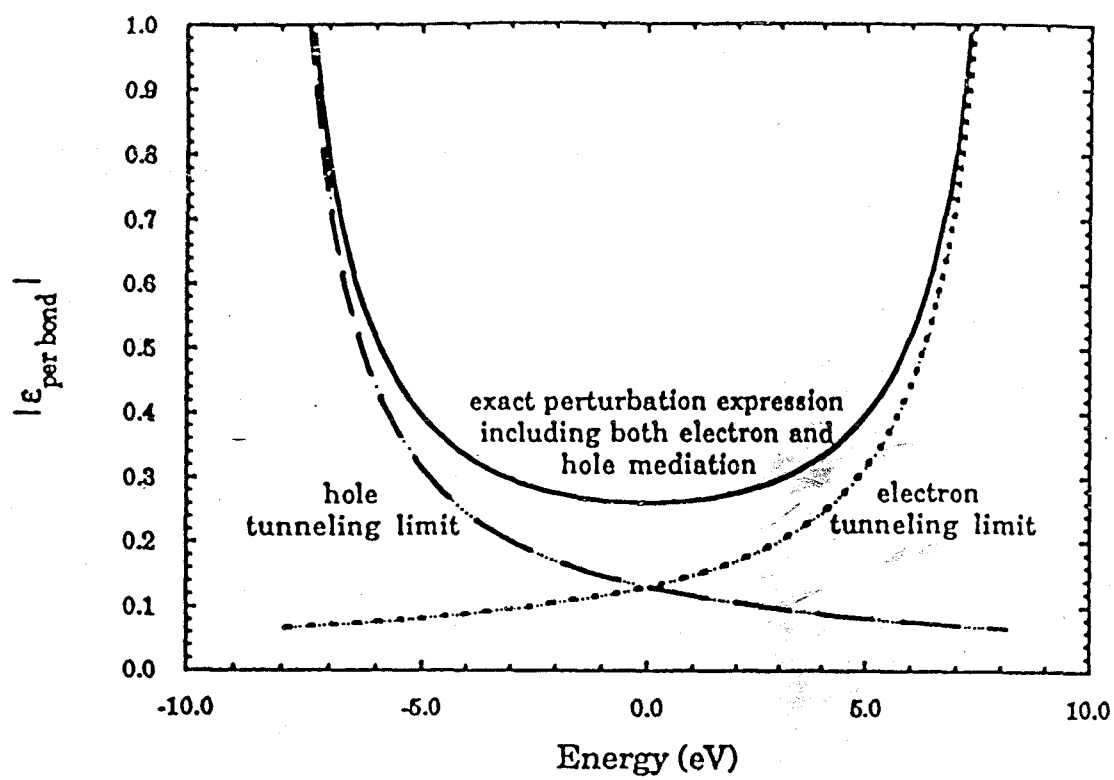


Figure 1.5. Inverted behavior is shown for D-A systems that are bridged by a steroid spacer.

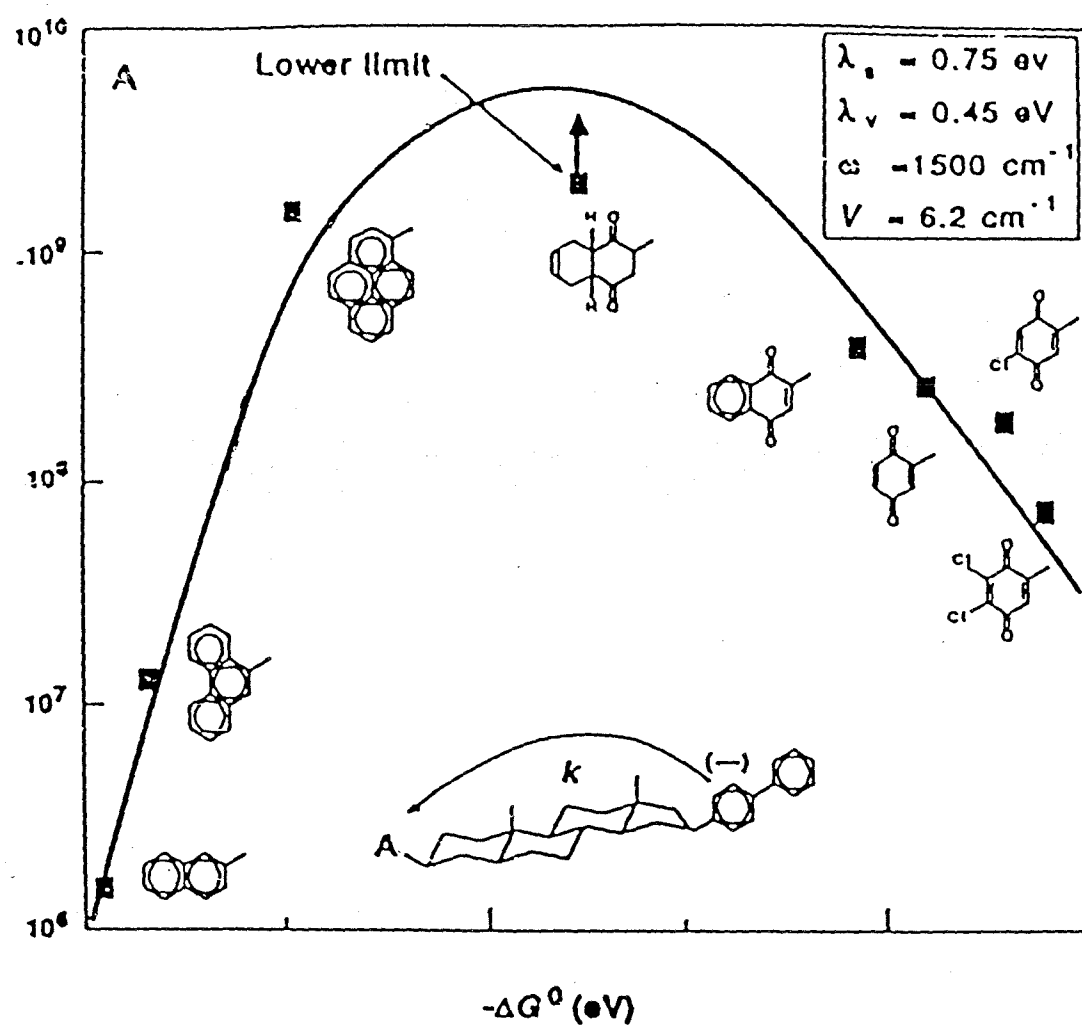


Figure 1.6. Plots of $\ln k$ as function of $-\Delta G^0$ for ruthenated cytochrome c derivatives. The best fit lines give an estimate for the reorganization energy (λ) of 1.2 eV.

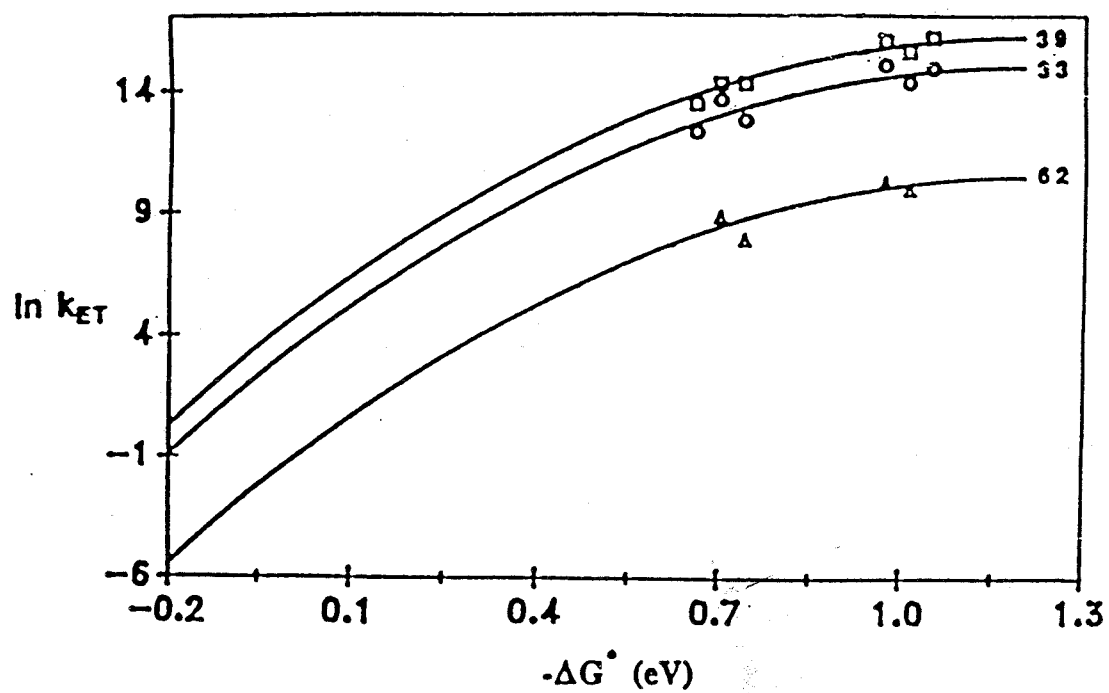


Figure 1.7. Positions of ruthenium attachment sites in cytochrome*c*. The dotted line in the H72 pathway indicates a through space jump.

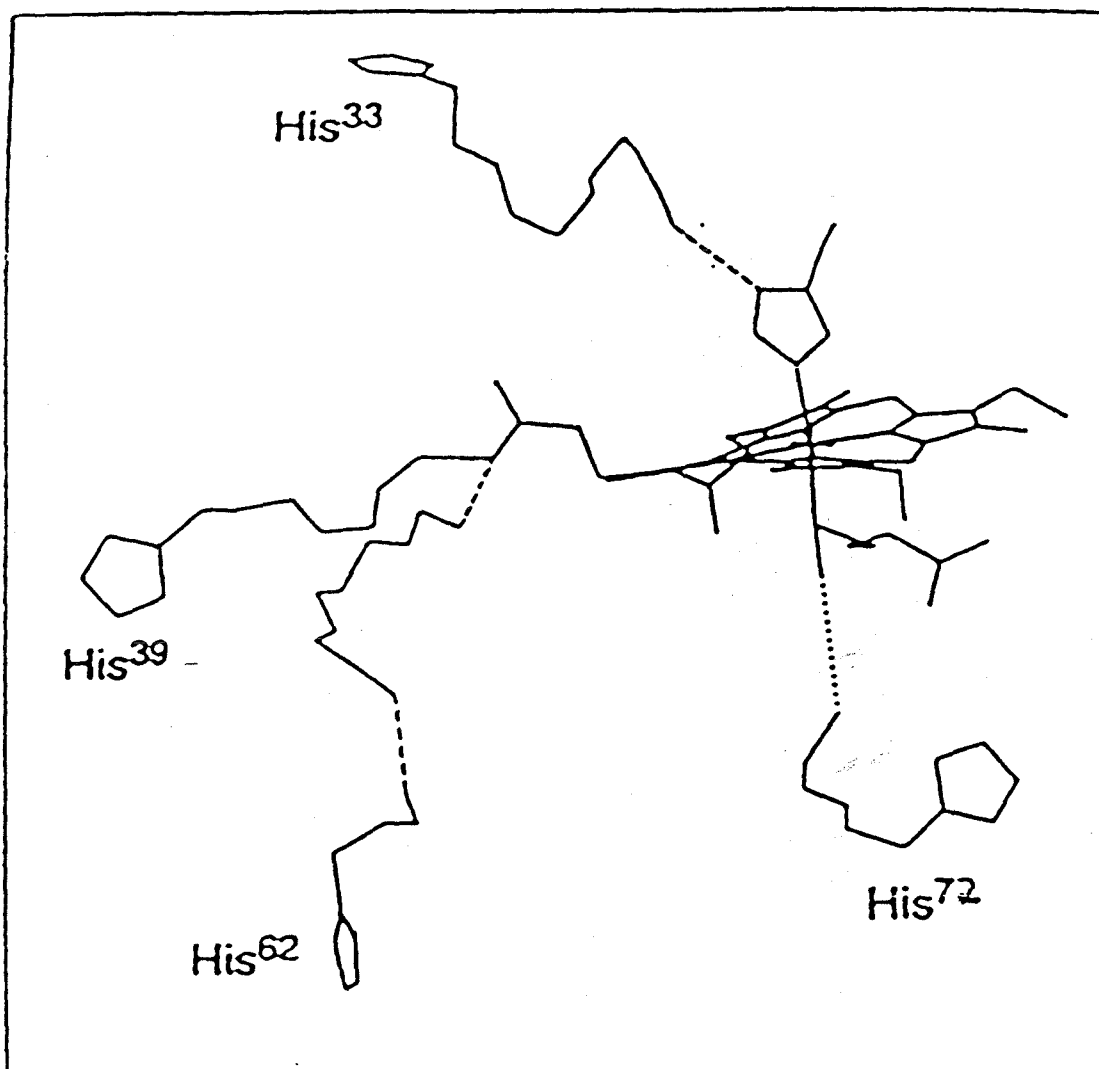
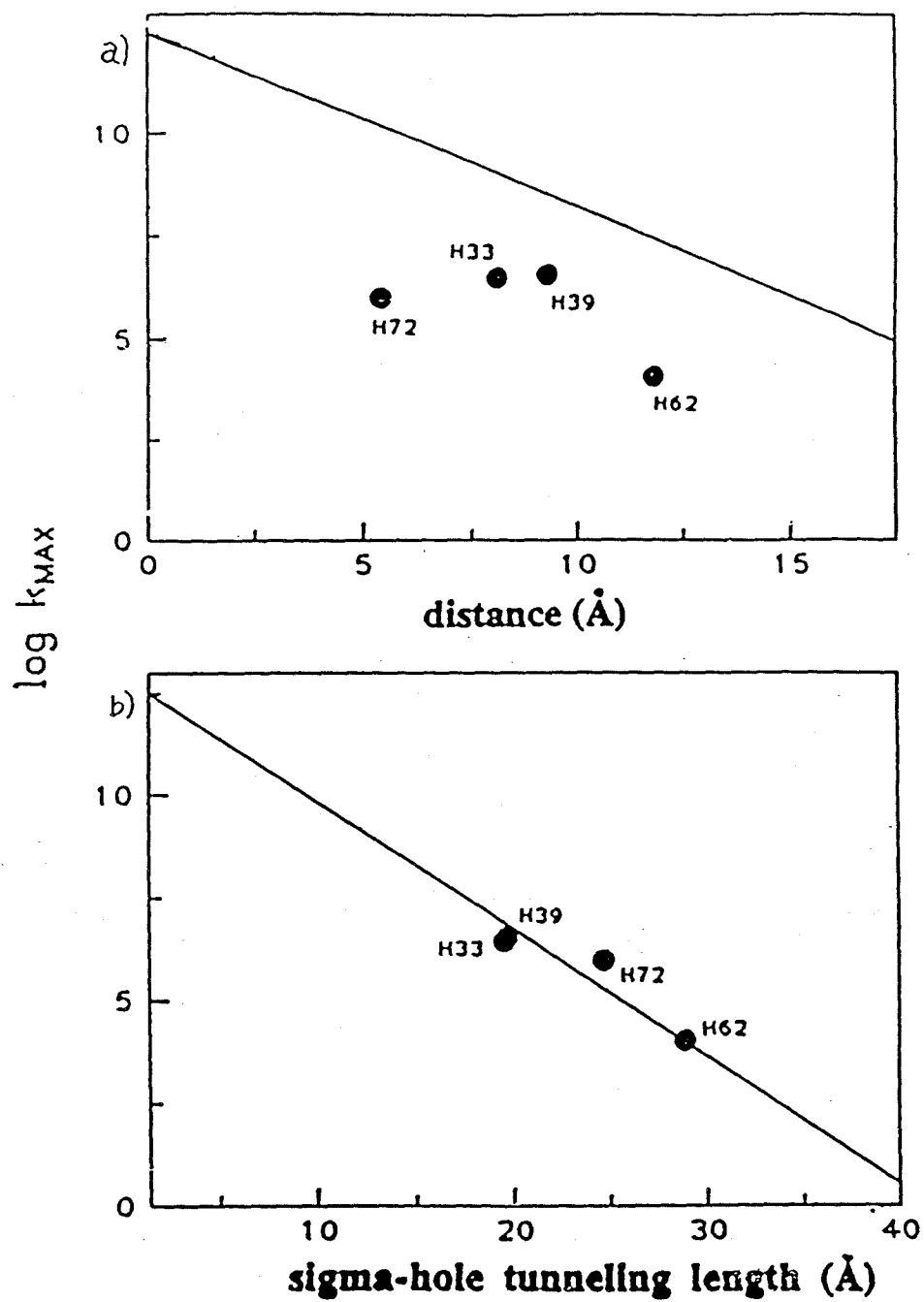


Figure 1.8. Log of measured electron transfer rates of ruthenium-modified cytochrome *c* mutants (H33, H39, H62, and H72) is plotted vs (a) distance and (b) effective pathway distance (measure of the calculated coupling). The solid line in (a) is derived from square barrier model with a β of 1.4/Å. Points lying on the solid line in (b) are correctly predicted by the B&O pathway model.



Chapter 2

Electronic Coupling through the M121-Strand in Azurin

ABSTRACT

Electronic coupling through a β -strand of *Pseudomonas aeruginosa* azurin has been investigated by measuring intramolecular electron-transfer (ET) rates in ruthenium-modified derivatives of the β -barrel blue copper protein. Surface histidines introduced on the Met121 β -strand by mutagenesis were modified with a $\text{Ru}(\text{bpy})_2(\text{im})^{2+}$ (bpy = 2, 2'-bipyridine; im = imidazole) complex. The Cu^+ to Ru^{3+} rate constants (H122, $7.1(4) \times 10^6$; H124, $2.2(2) \times 10^4$; H126, $1.3(6) \times 10^2 \text{ s}^{-1}$) yield a distance-decay constant of 1.1 \AA^{-1} , a value close to the decay of 1.00 \AA^{-1} predicted for electron tunneling through an idealized β -strand. Activationless ET rate constants in combination with a tunneling-pathway analysis of the structures of azurin and cytochrome *c* confirm that there is a generally efficient network for coupling the internal (native) redox center to the surface of each protein.

INTRODUCTION

Long-range protein electron-transfer (ET) reactions are key steps in many biological processes [1, 2]. According to theory [3], the rates of these ET reactions are proportional to an electronic coupling factor, $|H_{AB}|^2$, and a Franck-Condon factor ($F.C.$) that depends upon the difference between the reaction driving force (ΔG°) and the nuclear reorganization energy (λ). Reactions with $-\Delta G^\circ = \lambda$ exhibit maximum rates (k_{max}) that are limited by $|H_{AB}|^2$. The simplest description of H_{AB} treats the medium between donor and acceptor as a one-dimensional square tunneling barrier (1DSB). As such, the rate is predicted to drop exponentially with distance [4, 5]:

$$k_{ET} = \frac{2\pi}{h} |H_{AB}(R_0)|^2 \exp[-\beta(R - R_0)](F.C) \quad (1)$$

Dutton and coworkers have demonstrated that a 1DSB model for electron tunneling (eq 1), with an exponential distance decay constant β of 1.4 \AA^{-1} , provides a rough estimate of coupling strengths for many protein ET systems over a wide range of donor-acceptor separations [6, 7].

However, both theoretical and experimental investigations have shown that the structure of the intervening protein medium must be included to provide a semiquantitative description of the tunneling process [8-28]. Beratan and Onuchic have developed a tunneling-pathway model to describe electronic couplings between redox sites in a protein [9-12]. The protein medium between the two redox sites is decomposed into smaller subunits linked by covalent bonds, hydrogen bonds, or through-space jumps. Each link is assigned a coupling decay, and a structure-searching algorithm is employed to identify the optimum coupling pathway between redox sites. This pathway can be described in terms of an effective covalent tunneling length (σl) and ET rates should drop exponentially with (σl) rather than the direct distance between redox sites. If a coupling

decay of 0.6 per covalent link is assumed, then the exponential decay constant will be 0.73 \AA^{-1} [29-32].

A straightforward prediction of structure-dependent models is that different secondary structures will exhibit different coupling efficiencies per unit distance. The individual strands of β -sheets, for example, define nearly linear coupling pathways along the peptide backbone. This contrasts with α -helices in which the distance between the termini increases much more slowly (and non linearly) as the peptide lengthens. Hence, at any given donor-acceptor distance the coupling through a β -strand should be greater than that through an α -helix. This point is illustrated in Fig. 2.1, which compares σ_1 for an idealized β -strand and an idealized α -helix as functions of the terminal β -carbon separation (R_β). The tunneling length for a β -strand exhibits an excellent linear correlation with R_β , while σ_1 for the α -helix shows a poorer correlation, especially for $R_\beta \leq 10 \text{ \AA}$. As expected, the best linear fits to the σ_1/R_β plots yield different slopes: $1.72 \sigma_1/\text{\AA}/R_\beta$ for the α -helix and $1.37 \sigma_1/\text{\AA}/R_\beta$ for the β -strand. The near-linear σ_1/R_β correlations for the α -helix and β -strand structures imply that ET rates in these structures should exhibit an exponential dependence on direct distance. Interestingly, the predicted distance decay constants for these two structures (α -helix, 1.26 \AA^{-1} (18); β -strand, 1.00 \AA^{-1}) are substantially smaller than the empirical 1.4 \AA^{-1} decay constant for proteins.

We have probed the electronic coupling along a β -strand of *Pseudomonas aeruginosa* azurin by attaching $\text{Ru}(\text{bpy})_2(\text{im})^{2+}$ (bpy = 2,2'-bipyridine; im = imidazole) [33] to the surface histidine in each of the three site-directed mutants (H122, H124, H126) of this β -barrel blue copper protein (Fig. 2.2). Intramolecular ET rates from Cu^+ to $\text{Ru}(\text{bpy})_2(\text{im})(\text{His}^X)^{3+}$ ($X = 122, 124, 126$) were measured by laser flash transient spectroscopy.

MATERIALS AND METHODS

DNA

The introduction of surface histidines and the removal of the wild-type H83 residue was accomplished using the Muta-Gene mutagenesis kit purchased from Bio-Rad, which was developed by Kunkel [34, 35]. The manufacturer's protocols were followed precisely. The strategy used is only outlined briefly here. The previously synthesized azurin gene [36] was cloned into pTZ18U. Using helper phage M13K07 a single stranded uracil containing DNA was isolated and checked by sequencing. This DNA served as template for all mutagenesis reactions.

Since all mutants were double mutants at distant sites two mutagenic oligonucleotides were used simultaneously during the synthesis of the complementary strand. The double stranded product was then used to transform MV1190 by electroporation. After plating onto LB-AMP plates colonies were picked and grown up for DNA boiling minipreps. The DNA obtained was checked by sequencing. Successfully mutated DNA was cut with Bgl II and NdeI. The gene fragment was separated from the vector using a standard low-melt gel method. It was ligated into a T7-promotor expression system [37], a pET-3a vector (Novagene) that was cut with Bam HI and NdeI.

Expression and Purification

The pET-3a vector containing the respective mutated azurin genes was used to transform BL21(DE3) cells. Cells were grown at 37°C in 6l LB media in the absence of CuSO₄. At the end of logarithmic phase (OD≈1.0) the cells were induced with 0.6 mM isopropyl β-D-thiogalactopyranoside. The osmotic extrusion protocol [36] was used in a slightly modified version for the purification of the protein. The cells were harvested after induction for 5-10 hours by centrifugation at 5000 g for 15 min. The pellet was redissolved in 600 ml of a 20% sucrose solution containing 30 mM Tris pH=8.0 and 1mM

EDTA pH8.0. After 10 min the suspension was spun down again. The pellet was redissolved in 600 ml of a 1mM CuSO₄ solution. After shaking for one hour at room temperature, it was centrifuged again. The supernatant was acidified with a few drops of acidic acid to reach a pH of 4.5. The precipitate was separated by centrifugation at 5000 g. Incorporation of copper as judged by the appearance of the blue absorption band was greatly accelerated by heating the supernatant to 45 °C in the presence of 10 mM CuSO₄ for at least one hour. The holoprotein was subsequently purified by FPLC using a MonoS column. A 25 mM NaOAc solution (pH=4.5) was used as loading buffer. The mutant proteins were eluted with a gradient reaching a concentration of 300 mM NaOAc (same pH). The unmodified proteins were typically eluted with 12% of the elution buffer present.

Ruthenium Modification

For the ruthenation reaction a somewhat modified version of ref. [33] was used. A 0.1 mM solution of azurin was modified at 37°C using a 0.2 mM solution of Ru(bpy)₂(CO₃) in freshly prepared 300 mM NaHCO₃ buffer pH=7.4. Typically, the reaction was monitored by taking time points in 2 hour intervals. For these time points aliquots of the reaction mixture were taken and run through a PD10 gel filtration column equilibrated with 25 mM NaOAc pH=4.5. Completeness of the reaction was estimated by the ratio of the ruthenium absorption band to the intrinsic absorption band of the protein. The reaction was stopped by gel filtration using 25 mM NaOAc pH=4.5 when a 1:1 ratio of protein to ruthenium was reached. The extinction coefficients for protein (at 625 nm) and ruthenium (at 488) were taken as 5000 and 9000 M⁻¹L⁻¹ respectively. The products were purified by FPLC as described above. The fractions were analyzed by UV-vis spectroscopy. The correct ruthenium derivative showed an absorption maximum of 488 nm. This fraction was incubated in a 0.4 M imidazole solution pH 8.2 for at least 48

hours affording Ru(bpy)₂(im)(His)-azurin. Again, the products were cleaned by FPLC as described above.

Characterization of the Ruthenium-Modified Product

The ruthenium-modified product was characterized by UV-Vis absorption spectroscopy. The absorption maximum of the histidine-bound Ru(II)(bpy)₂(im) was 490 to 492 nm with a characteristic shoulder around 450 nm. Further characterization was accomplished by tryptic digestion, which required denaturation of the protein using trichloroacetic acid. The denatured protein was spun down in a Beckman microcentrifuge for 1 min. The pellet was washed with acetone, dried and then dissolved in 0.1 M NH₄HCO₃, pH 8.0 yielding a concentration of 2mg/ml. A 2mg/ml stock solution of trypsin was prepared in 1mM HCl. 20 µl of this solution were added to the protein solution. The reaction was carried out at 25°C. After 12 hours another 20 µl of the trypsin stock solution was added. After another 14 hours of reaction the reaction was quenched by flash-freezing using an ethanol-dry ice mixture and subsequently lyophilized. The lyophilized pellet was then redissolved in a 0.1% aqueous TFA solution and loaded onto a reversed-phase PepRC HR 5/5 column (Pharmacia). A 0.1% solution of TFA in acetonitrile was used as elution buffer (buffer B) using a linear gradient from 0 to 40% buffer B. The peptides were sequenced with an ABI 473A sequencer.

Rate Measurements

Two methods for rate determination were used, flash-quench technique as well as photoinduced technique [33] (Fig.2.3).

In the flash-quench experiment the Ru(II)(bpy)₂(im) bound to a surface histidine is excited by a 480 nm laser pulse. The excited state is quenched by a 5 μ M solution of Ru(III)(NH₃)₆. The resulting Ru³⁺(bpy)₂(im) is then re-reduced by intramolecular ET from the reduced copper center with a rate of k_{ET}^b . In the case of H126, the rate for the intramolecular ET (k_{ET}^b) had to compete with the back reaction in which the photochemically generated Ru(NH₃)₆ rereduces the surface bound ruthenium complex. Therefore rates were also measured using the irreversible quencher Co(NH₃)₅Cl₃. Due to concentration dependence of the ET rate for this mutant, rates were obtained by extrapolation to infinite dilution of protein.

The photoinduced electron transfer involves an initial charge separation and the reverse charge recombination (Fig 2.3). The charge separation is induced by laser excitation of the surface bound Ru(II)(bpy)₂(im). In the absence of an external quencher the excited state is partially being quenched by the copper site with a rate of k_{ET}^f . Then the system returns into the original state with a rate of k_{ET}^b . The kinetics for the electron transfer steps were monitored by the change of absorption at independent wavelengths. The reduction and oxidation of the copper center was monitored by the changes in absorbance at 628nm. The oxidation and reduction of the surface Ru was monitored at 500, 432, and 310 nm.

Temperature Dependence of Electron Transfer Rates

The photoinduced scheme was used to measure the temperature dependence for the H124 mutant. Rates were measured at 625 and 500 nm. The temperature range was varied from 30°C to 47°C. A minimum of 30 min of temperature equilibration was allowed before every measurement.

Computer Modeling of the Mutant Structures

All distances were obtained by molecular modeling using the program ENZYMIX [38]. Starting from the X-ray structure [26, 27, 39] initial guesses of the histidine-modified azurin mutants were generated by manually docking a Ru(bpy)₂(im) complex onto the β -carbon of the respective mutated residue. This was done using the program MOLARIS such that unfavorable van-der-Waals contacts were avoided. The β -carbon was connected to the imidazole and resulted in a histidine side chain. This structure was relaxed by 2 ps of low temperature molecular dynamics (MD) at 10 K to relieve some strain. The resulting structures were then subjected to MD simulations at 300 K for 60 to 100 ps using the program ENZYMIX. The stepsize for the simulations was 2 fs, the non-bonded interactions were updated every 20 steps. Atoms outside a 12 Å radius around the ruthenium ion were constrained to their crystallographic positions. Water molecules within 14 Å of the ruthenium ion were modeled explicitly. From 14 to 18 Å water molecules were represented by Langevin-type dipoles. Only the copper center and the ruthenium ion were assigned net charges (+2 in both cases). The metal-to-metal distance between the ruthenium and the copper was obtained by averaging the distances in snapshots taken in 1 ps intervals. As a control the final converged structures were perturbed by changing the metal-to-metal distances by as much as 4 Å through the applications of harmonic distance constraints. When the simulations were continued without the constraints, distances moved back to the original average values.

RESULTS

DNA

The mutagenesis reactions were done according to the Bio-Rad protocols. The use of two separate mutagenic oligonucleotides decreased the yields to approximately 20 to 30%. The efficiency of the subsequent ligation into the pET-3a vector using an in-gel ligation protocol was rather poor. Significantly better yields were obtained using a gelase protocol.

Expression and Purification

All mutants were expressed equally well with average yields ranging from 30 to 50 mg/liter of bacterial culture. During the reconstitution of azurin with copper it was found that some fraction of the copper was incorporated into the protein within only a few minutes at room temperature in 25 mM NaOAc buffer pH=4.5. However, copper uptake was not complete as the blue color of azurin still increased after days. Heating the reconstitution reaction to 45 °C significantly accelerated this reaction and after approximately one hour the copper uptake was complete under those reactions. It appeared that copper uptake into apo-azurin is very quickly, even at room temperature. However, it is likely that some of the expressed protein was bound to zinc, which has been documented in the literature [40]. We believe that zinc dissociation is highly increased at 45 °C. Therefore, the copper uptake reaction was most likely accelerated under those conditions.

Ruthenium Modification

The ruthenium derivative Ru(bpy)₂(H₂O)(His)-azurin was purified using a MonoS ion exchange column (Pharmacia) and eluted at approximately 12% buffer B for the H122 mutant and at 18 to 20 % buffer B for the H124 and the H126 mutant. Typical yields at

this step were 80, 40, and 25 % for the H122, H124, and H126 mutant respectively. The subsequent imidazole reaction had a yield of at least 80 % in all cases.

Rate Measurements

Figs. 2.4 and 2.5 show representative time-resolved absorption kinetics for the H122, and the H124 mutant. Rates were obtained from both flash-quench and photoinduced ET. The slow ET rate in the H126 derivative led to a number of complications. With $\text{Ru}(\text{NH}_3)_6^{3+}$ as a quencher, $\text{Cu}^+ \rightarrow \text{Ru}^{3+}$ ET was in competition with $\text{Ru}(\text{NH}_3)_6^{3+} \rightarrow \text{Ru}^{3+}$ charge recombination. Consequently, kinetics were also measured using the irreversible quencher $\text{Co}(\text{NH}_3)_5\text{Cl}^{2+}$; consistent results were obtained with the two quenchers. At the protein concentrations used in these experiments (4-60 μM), bimolecular $\text{Cu}^+ \rightarrow \text{Ru}^{3+}$ ET in the H126 derivative occurred on the same time scale as intramolecular $\text{Cu}^+ \rightarrow \text{Ru}^{3+}$ ET; the intramolecular ET rate was extracted from extrapolation of the observed ET rates to infinite dilution (Fig.2.7). The rate constants for the H122 , H124 , and H126 mutants were $7.1(4) \times 10^6 \text{s}^{-1}$, $2.2(2) \times 10^4 \text{s}^{-1}$, and $1.3(6) \times 10^2 \text{s}^{-1}$, respectively.

Temperature Dependence of Electron Transfer

The temperature dependence of ET for the H124 mutant is shown in Fig. 2.8. It can be seen that temperature has only a very small effect on the ET rate as it decreases only slightly with increased temperature.

Characterization of the Ruthenium-Modified Product

As an example of the characterization of the ruthenium-modified protein the sequencing result for the H124 tryptic digestion fragment is shown in Fig.2.9. It can be seen that the digestion must have occurred after K122 as is expected for Trypsin since the sequence of the peptide corresponds to residues 123 through 128 in azurin. Interestingly, the intensity

with which the H124 was found is decreased compared to the other amino acids consistent with a ruthenium modification.

DISCUSSION

The driving force for the ET reactions measured in this chapter are estimated at $-\Delta G^\circ = 0.75$ eV, which is nearly equal to the estimated reorganization energy for this system ($\lambda \approx 0.8$ eV) [28]. The observed rates should therefore be within 5% of the activationless rates (k_{max}). This is consistent with the small temperature dependence of the ET rate for the H124 mutant. A plot of $\log(k_{max})$ vs. R_M (R_M = metal-metal separation) for Ru-modified azurin is nearly linear (Fig. 2.10); the best fit has a distance decay constant of 1.10 \AA^{-1} and a close contact rate ($R_M = 3 \text{ \AA}$) of 10^{13} s^{-1} . It is encouraging that the decay constant is quite close to that predicted for coupling along a β -strand. The deviation from the β -strand line can be attributed to the fact that the redox sites do not lie on the β -carbon atoms, but instead are separated from them by the distances from the β -carbons to the metals.

An analysis of the tunneling pathways in azurin requires consideration of the unusual coordination environment of the Cu center. X-ray structure determinations of the azurin active site reveal a Cu atom with approximate trigonal bipyramidal coordination [39]. The equatorial plane is comprised of three normal ligands, His46, His117, and Cys112. In the axial positions are the carbonyl oxygen of Gly45 at 3.22 \AA and the sulfur atom of Met121 at 3.22 \AA . These long metal-axial ligand distances complicate calculations of tunneling pathways because it is not clear whether these interactions should be treated as normal covalent bonds or as through-space gaps. For the Ru(His^X)-azurin system (X = 122, 124, 126), the coupling involves Met121; the one-bond rate estimated from a $\log k_{max}/\sigma l$ correlation is 10^{13} s^{-1} when the Cu-Met121 interaction is treated as a space

jump. The pathway model predicts that the coupling across space gaps drops off as an exponential function of distance (decay constant of 3.4 \AA^{-1}) [9-12]. For the Cu-Met121 interaction, the gap is therefore taken as the difference between the observed Cu-S separation and a normal Cu-S bond distance ($\approx 2.3 \text{ \AA}$). A more detailed discussion can be found in chapter 3. If the Cu-Met121 interaction is assumed to be a normal covalent bond, the one-bond limit is only 10^{12} s^{-1} . These values of the one-bond rates suggest that the Met121 sulfur is weakly coupled to the Cu atom, in close agreement with the electronic structure calculations of Lowery and Solomon [41].

Using the tunneling-pathway model, it is possible to compare the results obtained with Ru-modified azurin to those from other Ru-modified proteins with similar donor and acceptor redox potentials. ET rates have been measured in eight different $\text{Ru}(\text{bpy})_2(\text{im})(\text{His}^X)$ -modified cytochromes *c* [28] but, unlike Ru-modified azurin, there is no uniform secondary-structure type separating the heme and the six Ru-binding sites. Nevertheless, a plot of σl vs. R_M (Fig. 2.11) reveals that a single straight line (slope = $1.45 \text{ } \sigma l \text{ -}\text{\AA}/R_M\text{-}\text{\AA}$) adequately describes all but two of the Ru binding sites ($\text{Ru}(\text{His}72)\text{-cyt } c$; $\text{Ru}(\text{His}62)\text{-cyt } c$). We would expect, then, that $\log k_{max}$ should vary linearly with R_M ($\beta = 1.06 \text{ \AA}^{-1}$) for all but the His72 and His62 data points, an expectation that is largely borne out by the results (Fig. 2.11). In this analysis, we have used the tunneling-pathway model to identify a family of Ru-binding sites on azurin and cytochrome *c* in which the electronic coupling can be described by an *effective* square-tunneling barrier. In view of the disparate coupling efficiencies of common protein secondary structures (Fig. 2.1), it is likely that families of binding sites on these (or other) proteins can be found that exhibit distinctly different effective tunneling-barrier heights (β).

Two key elements determine the distance dependence of electronic coupling in proteins, the directness of the coupling pathway and the energy of the tunneling electron. The former depends upon the protein structure and the positions of the redox sites. In the

tunneling-pathway model, the tunneling energy defines the covalent-bond coupling decay (E_C); the value of 0.6 derives from analyses of ET rates in model systems [9-12, 28]. Changes in E_C will not affect σ/R_M correlations, since E_C appears as a prefactor in the expressions for hydrogen-bond and space-gap coupling decays. The magnitude of E_C does, however, determine the predicted slopes in $\log k_{max}/\text{distance}$ plots. The close agreement between the observed and predicted decay constants for azurin and cytochrome *c* suggests that an E_C value of 0.6 is appropriate for a high-potential oxidant such as $\text{Ru}(\text{bpy})_2(\text{im})(\text{His}^X)^{3+}$ ($E^\circ = 1.0 \text{ V vs. NHE}$). In the systems analyzed by Dutton [6, 7], the redox couples spanned a fairly wide range, but they included many lower potential oxidants. Even accounting for the generally less favorable tunneling energies, it is surprising that Dutton's average protein distance decay (1.4 \AA^{-1}) is so much greater than that found for the electron-transfer proteins azurin and cytochrome *c*. We suspect that part of the steeper decay reflects the inhomogeneity of typical proteins in which coupling must be mediated by through-space gaps rather than idealized secondary-structure elements. It also should be noted that donor-acceptor electronic coupling decays much more rapidly with distance in Ru-modified myoglobin, an oxygen-binding protein, than in azurin or cytochrome *c* [29-32].

BIBLIOGRAPHY

1. Bertini, I., Gray, H.B., Lippard, S.J., and Valentine, J.S.: Bioinorganic Chemistry, Mill Valley, CA: University Science Books; 1994.
2. Sigel, H., and Sigel, A.: Metal Ions in Biological Systems, New York: Dekker; 1991.
3. Marcus, R.A., and Sutin, N., *Biochim Biophys. Acta* **811**, 265 (1985)
4. Jortner, J., *J. Chem. Phys.* **64**, 4860 (1976)
5. Hopfield, J.J., *Proc. Natl. Acad. Sci.* **71**, 3640 (1974)
6. Farid, R.S., Moser, C.C., and Dutton, P.L., *Curr. Opin. Struct. Biol.* **3**, 225 (1993)
7. Moser, C.C., Keske, J.M., Warncke, K., Farid, R.S., and Dutton, P.L., *Nature* **355**, 796 (1992)
8. Friesner, R.A., *Structure* **2**, 339 (1994)
9. Onuchic, J.N., and Beratan, D.N., *J. Chem. Phys.* **92**, 722 (1990)
10. Onuchic, J.N., Andrade, P.C.P., and Beratan, D.N., *J. Chem. Phys.* **95**, 1131 (1991)
11. Onuchic, J.N., Beratan, D.N., Winkler, J.R., and Gray, H.B., *Annu. Rev. Biophys. Biomol. Struct.* **21**, 349 (1992)
12. Skourtis, S.S., Regan, J.J., and Onuchic, J.N., *J. Phys. Chem.* **98**, 3379 (1994)
13. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **94**, 8430 (1990)
14. Siddarth, P., and Marcus, R.A., *J. Chem. Phys.* **94**, 2985 (1990)
15. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **96**, 3213 (1992)
16. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **97**, 2400 (1993)
17. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **97**, 13078 (1993)
18. Stuchebrukhov, A.A., and Marcus, R.A., *J. Phys. Chem.* in press
19. Gruschus, J.M., and Kuki, A., *J. Phys. Chem.* **97**, 5581 (1993)
20. Broo, A., and Larsson, S., *J. Phys. Chem* **97**, 5581 (1993)
21. Christensen, H.E.M., Conrad, L.S., Mikkelsen, K.V., and Ulstrup, J., *J. Phys. Chem* **96**, 4451 (1992)
22. Evenson, J.M., and Karplus, M., *Science* **262**, 1247 (1993)

23. Farver, O., and Pecht, I., *J. Am. Chem. Soc.* **114**, 5764 (1992)
24. Farver, O., Skov, L.K., Pascher, T., Karlsson, G., Nordling, M., Lundberg, L.G., Vänngård, T., and Pecht, I., *Biochemistry* **32**, 7317 (1993)
25. Canters, G.W., and van de Kamp, M., *Curr. Opin. Struct. Biol.* **2**, 859 (1992)
26. Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G.W.J., *J. Mol. Biol.* **221**, 765 (1991)
27. Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G.W.J., *J. Mol. Biol.* **218**, 427 (1991)
28. Karpishin, T.B., Grinstaff, M.W., Komar-Panicucci, S., McLendon, G., and Gray, H.B., *Structure* **2**, 415 (1994)
29. Casimiro, D.R., Wong, L.-L., Colon, J.L., Zewert, T.E., Richards, J.H., Chang, I.-J., Winkler, J.R., and Gray, H.B., *J. Am. Chem. Soc.* **115**, 1485 (1993)
30. Casimiro, D.R., Richards, J.R., Winkler, J.R., and Gray, H.B., *J. Phys. Chem* **97**, 13073 (1993)
31. Wuttke, D.S., Bjerrum, M.J., Winkler, J., and Gray, H.B., *Science* **256**, 1007 (1992)
32. Wuttke, D.S., Bjerrum, M.J., Chang, I.-J., Winkler, J.R., and Gray, H.B., *Biochim. Biophys. Acta* **168**, 1101 (1992)
33. Chang, I.-J., Gray, H.B., and Winkler, J.R., *J. Am. Chem. Soc.* **113**, 7056 (1991)
34. Kunkel, T.A., Bebenek, K., and McClary, J., *Methods Enzymol.* **204**, 125 (1991)
35. Kunkel, T.A., *Proc. Natl. Acad. Sci.* **82**, 488 (1985)
36. Chang, T.K., Iverson, S.A., Rodrigues, C.G., Kiser, C.N., Lew, A.Y., Germanas, J.P., and Richards, J.H., *Proc. Natl. Acad. Sci.* **88**, 1325 (1991)
37. Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorf, J.W.: *Methods in Enzymology*, San Diego: Academic Press; 1990.
38. Lee, F.S., Chu, Z.T., and Warshel, A., *J. Comp. Chem.* **14**, 161 (1993)
39. Adman, E.T., Jensen, L.H., *Isr. J. Chem.* **21**, 8 (1981)
40. Nar, H., Huber, R., Messerschmidt, A., Filippou, A.C., Barth, M., Jaquinod, M., van de Kamp, M., and Canters, G.W., *Eur. J. Bioch* **205**, 1123 (1992)
41. Lowery, M.D., and Solomon, E.I., *Inorg. Chim. Acta* **200**, 233 (1992)

Figure 2.1. Plots of tunneling path lengths vs. β -carbon separation for idealized peptide secondary structures: α -helix neglecting hydrogen-bond interactions (o); α -helix including hydrogen-bond interactions (\bullet); β -strand (\bullet). Solid lines are best fits constrained to an intercept at the origin.

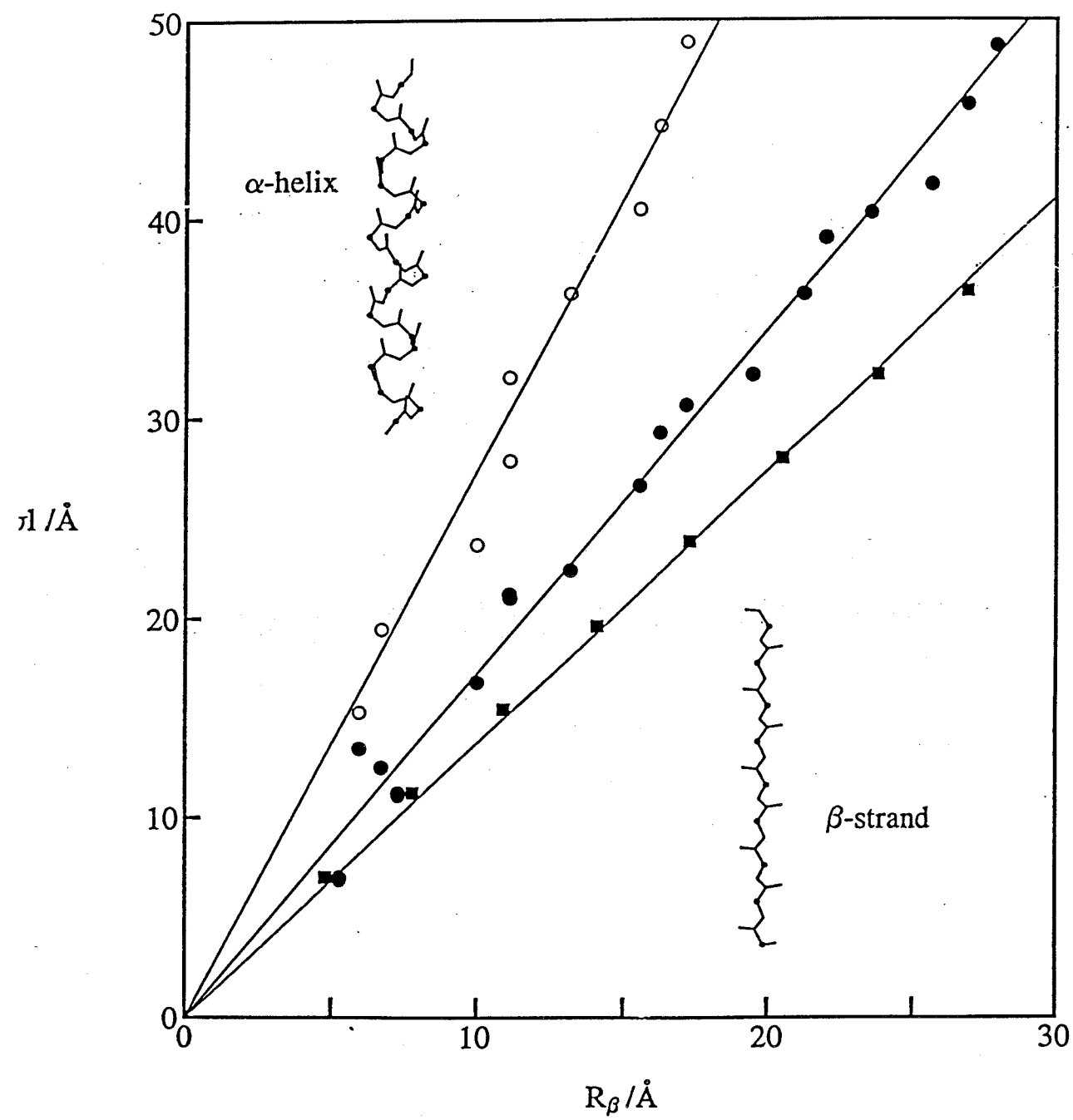
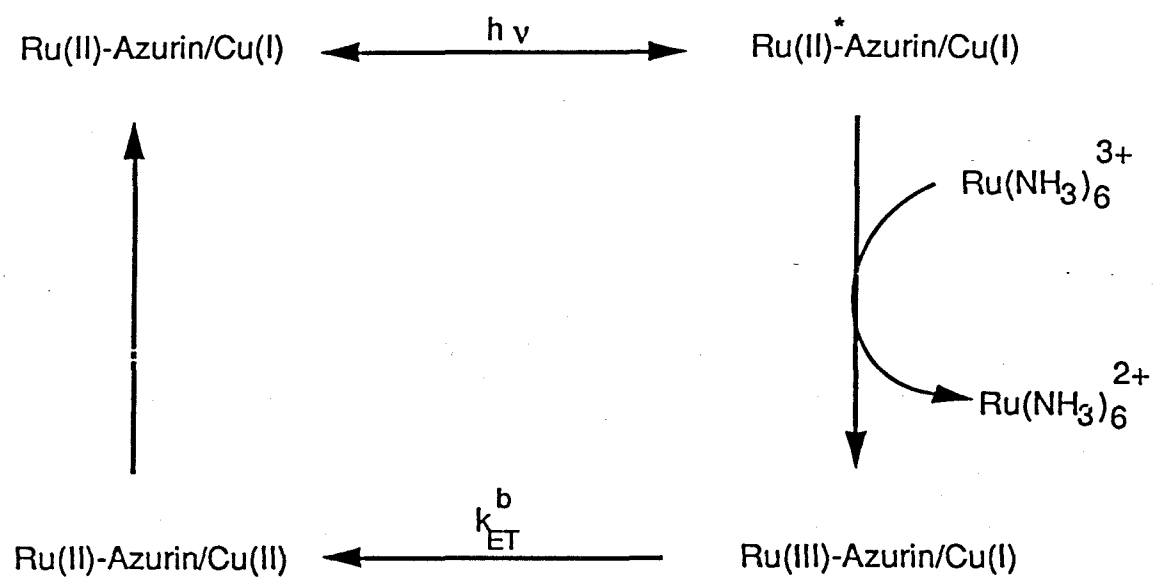


Figure 2.2. Backbone structure of azurin illustrating the location of the three surface His residues on the β -strand extending from the M121 ligand of the Cu atom.



Figure 2.3. Schematic drawing of the flash-quench and photoinduced techniques. Laser excitation is denoted by $h\nu$. k_{ET}^f and k_{ET}^b stand for forward rate and backward ET rate respectively.

Flash Quench



Photoinduced

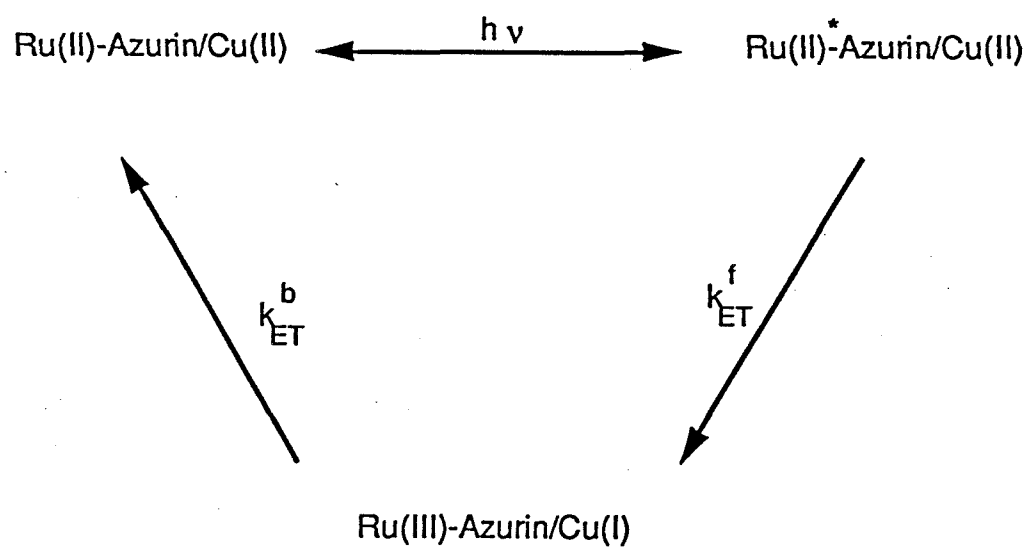
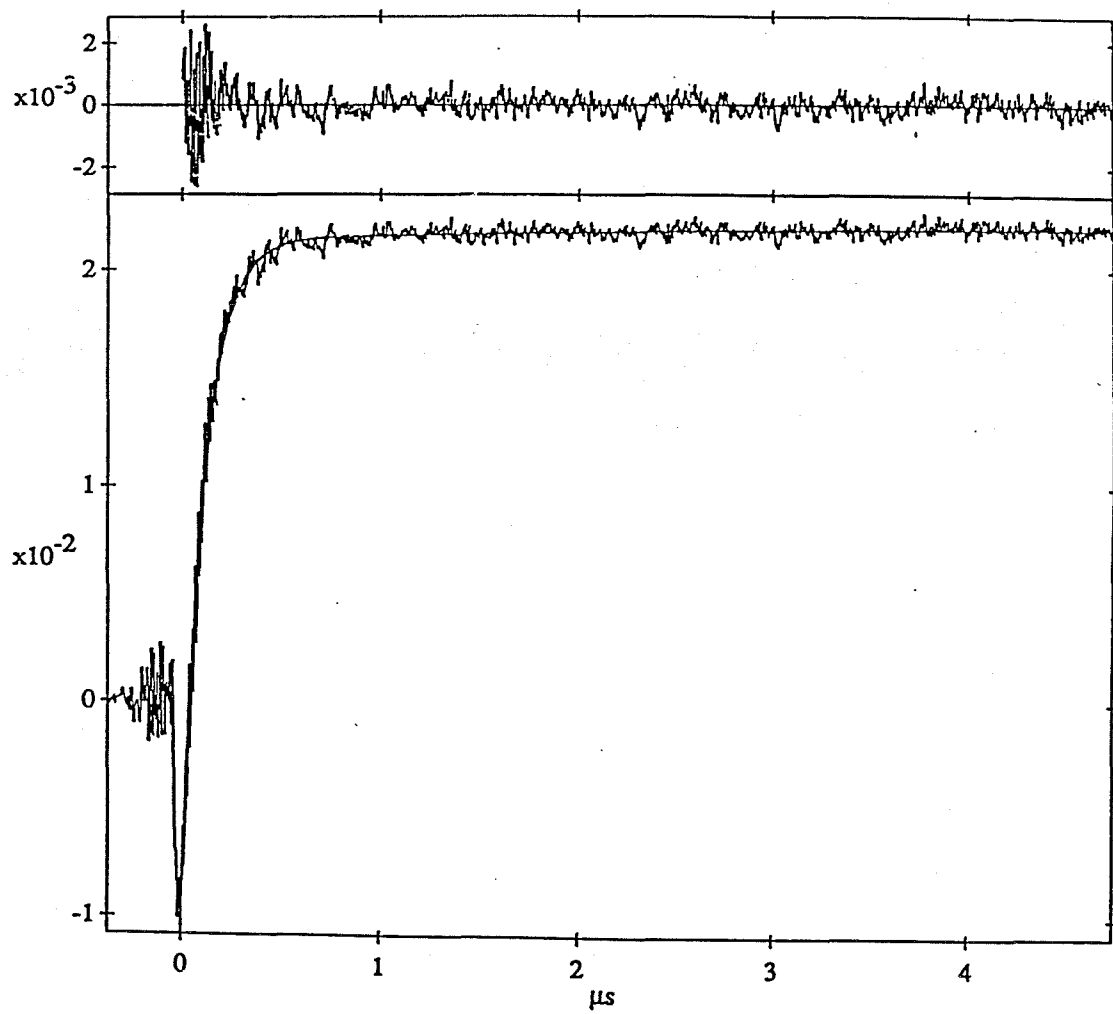
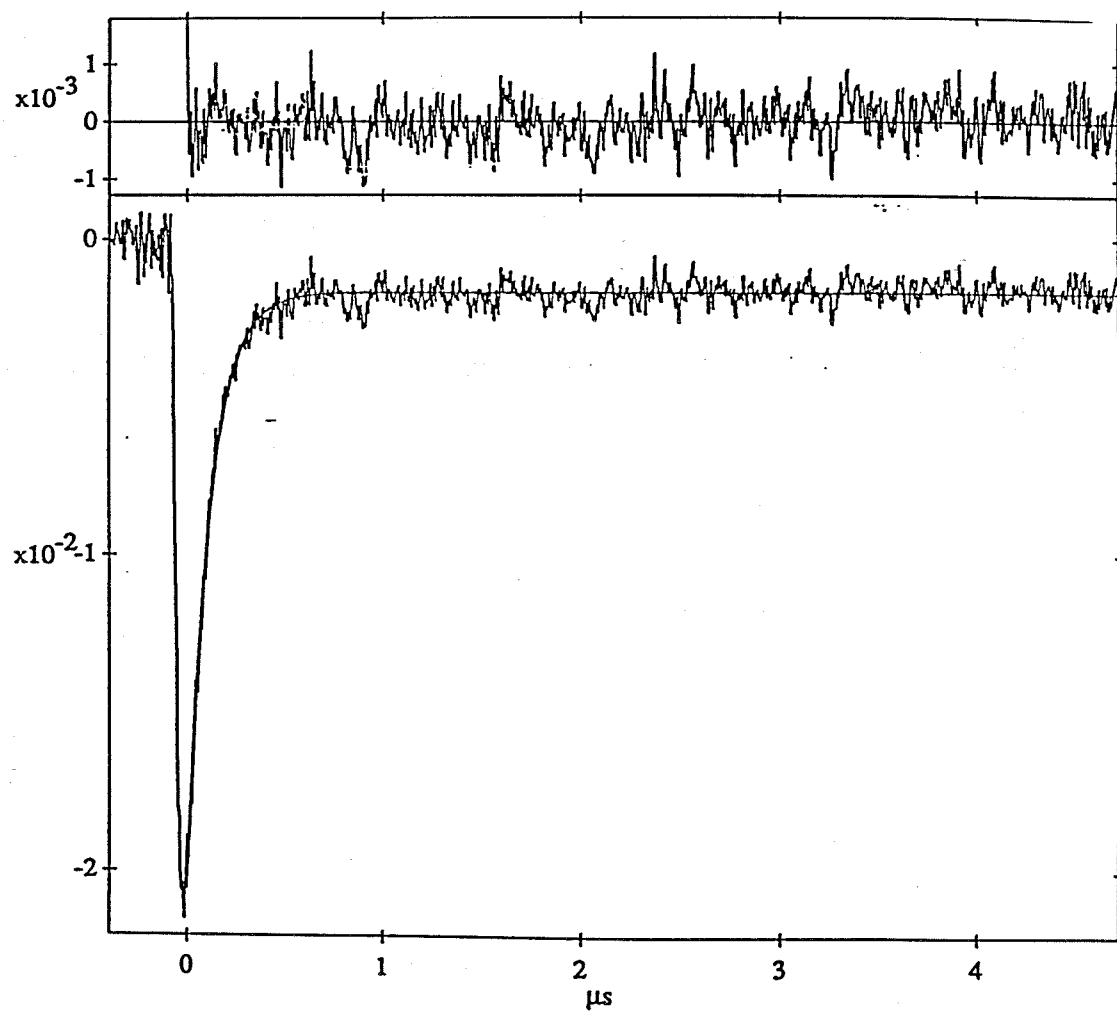


Figure 2.4. Representative kinetics are from flash-quench and photoinduced experiments with the H122 mutant. (a) The buildup of Cu^{2+} is monitored by the change in absorbance at 628 nm in a flash-quench experiment. The first short-lived bleach is due to the excited state decay. The subsequent increase in optical density is due to the buildup of Cu^{2+} . (b) Flash-quench experiment, in which the changes in absorption were monitored at 430 nm; at this wavelength hardly any signal from the excited state interferes with the ET. The bleach is due to the formation of Ru^{3+} . The recovery phase (Ru^{3+} to Ru^{2+}) is used to determine the ET rate (k_{ET}^b). (c) Changes in absorption in a photoinduced experiment monitored at 625 nm. The bleach is due to the formation of Cu^+ and the excited state emission. All data gave the same ET rate constant ($\text{Cu}^+ \rightarrow \text{Ru}^{3+}$) of $7 \times 10^6 \text{ s}^{-1}$.





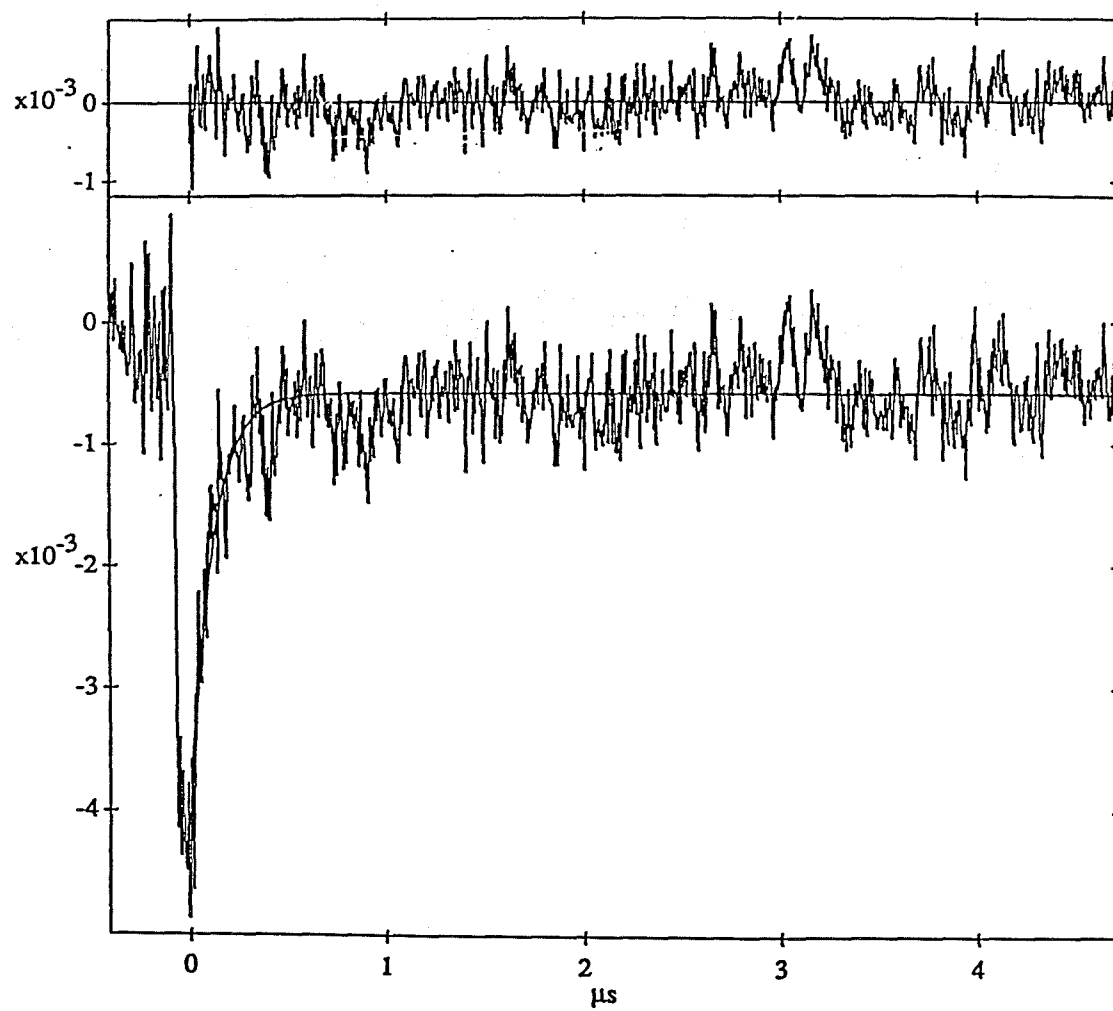
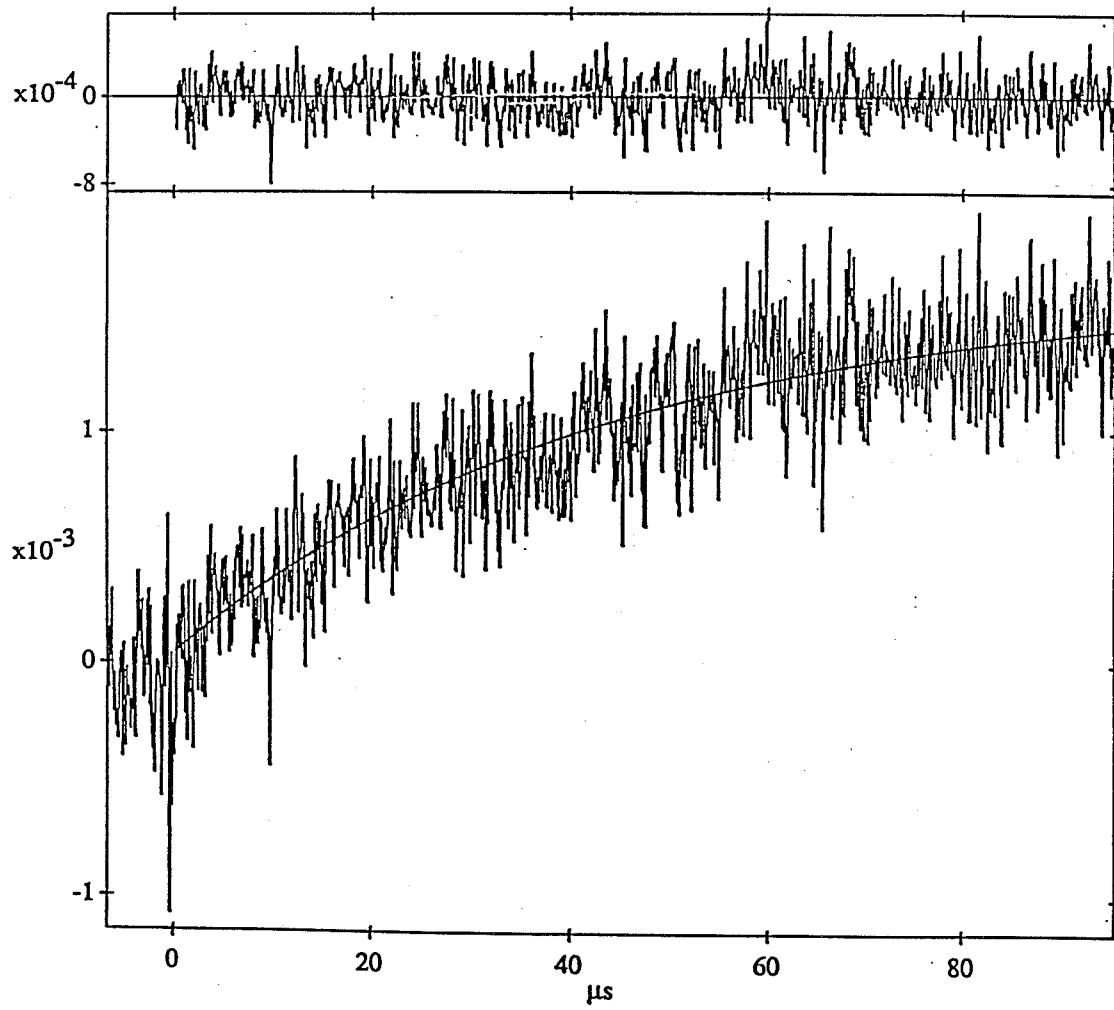
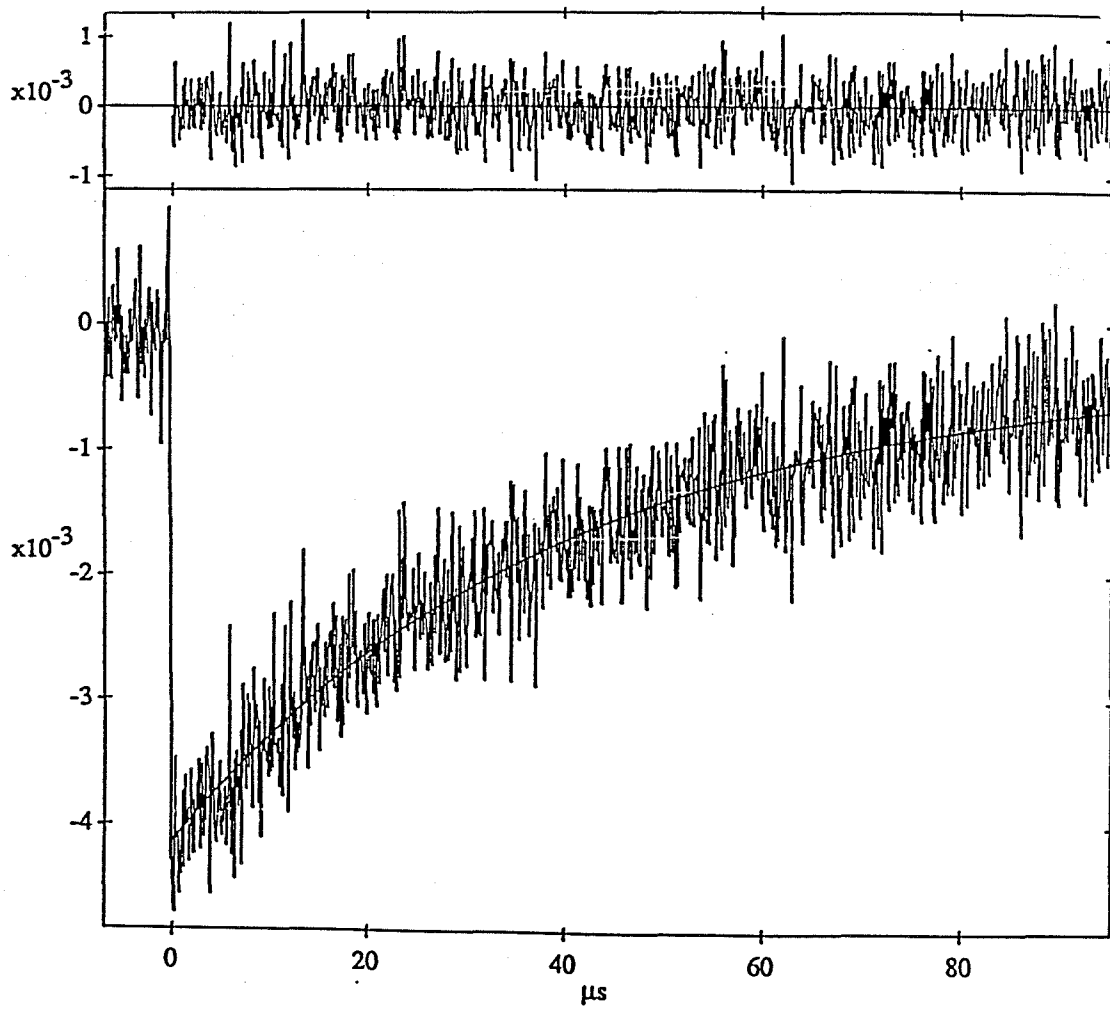


Figure 2.5. Representative kinetics are from (a) flash-quench and (b, c) photoinduced experiments with the H124 mutant. (a) The buildup of Cu^{2+} was monitored by the change in absorbance at 625 nm. (b) The change in absorption is monitored at 625 nm. The original bleach mainly corresponds to the formation of Cu^+ . The subsequent oxidation gives Cu^{2+} leads to an increase in the absorption. (c) The changes in absorption are monitored at 430 nm. All traces give the same rate constant for the $\text{Cu}^{1+} \rightarrow \text{Ru}^{3+}$ ET reaction of $2.2 \times 10^4 \text{ s}^{-1}$.





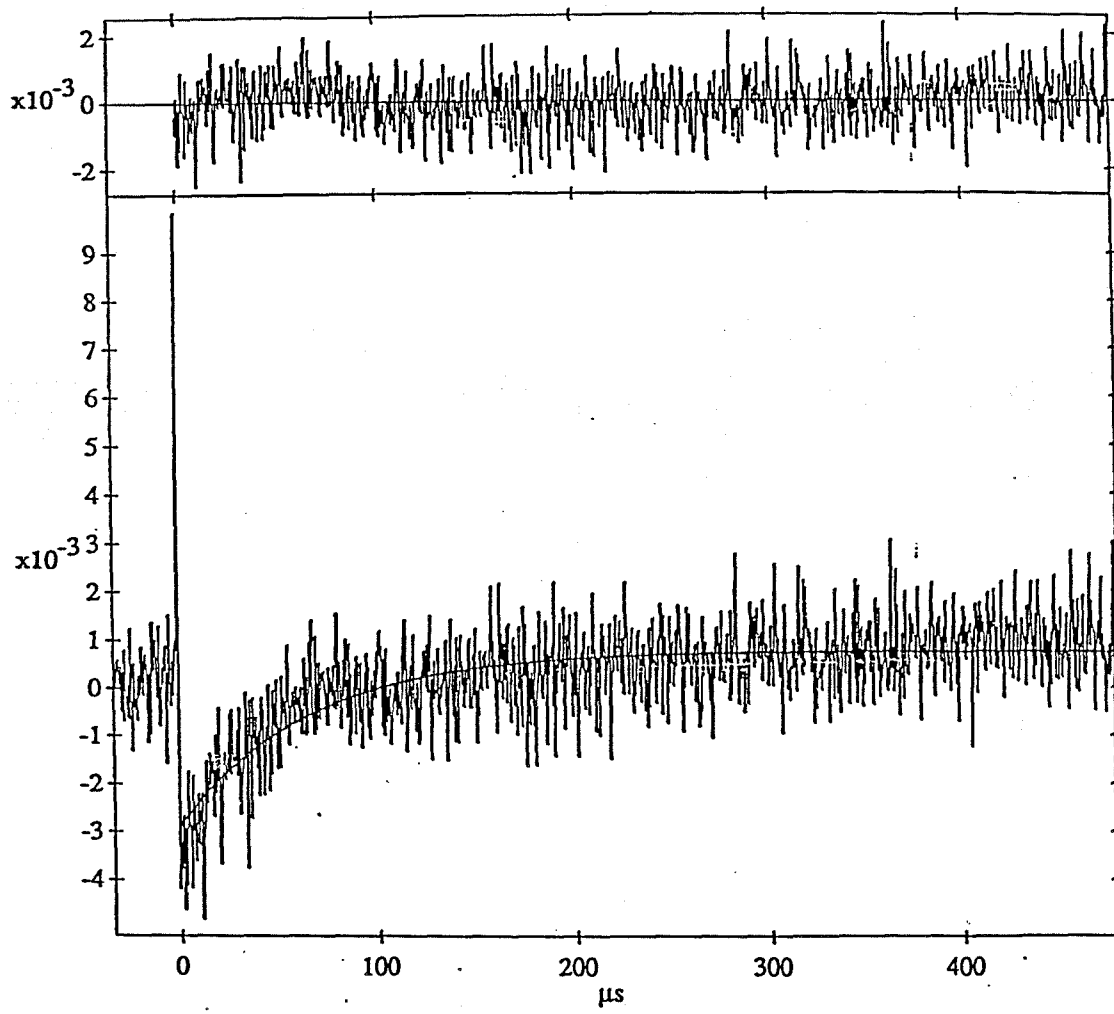
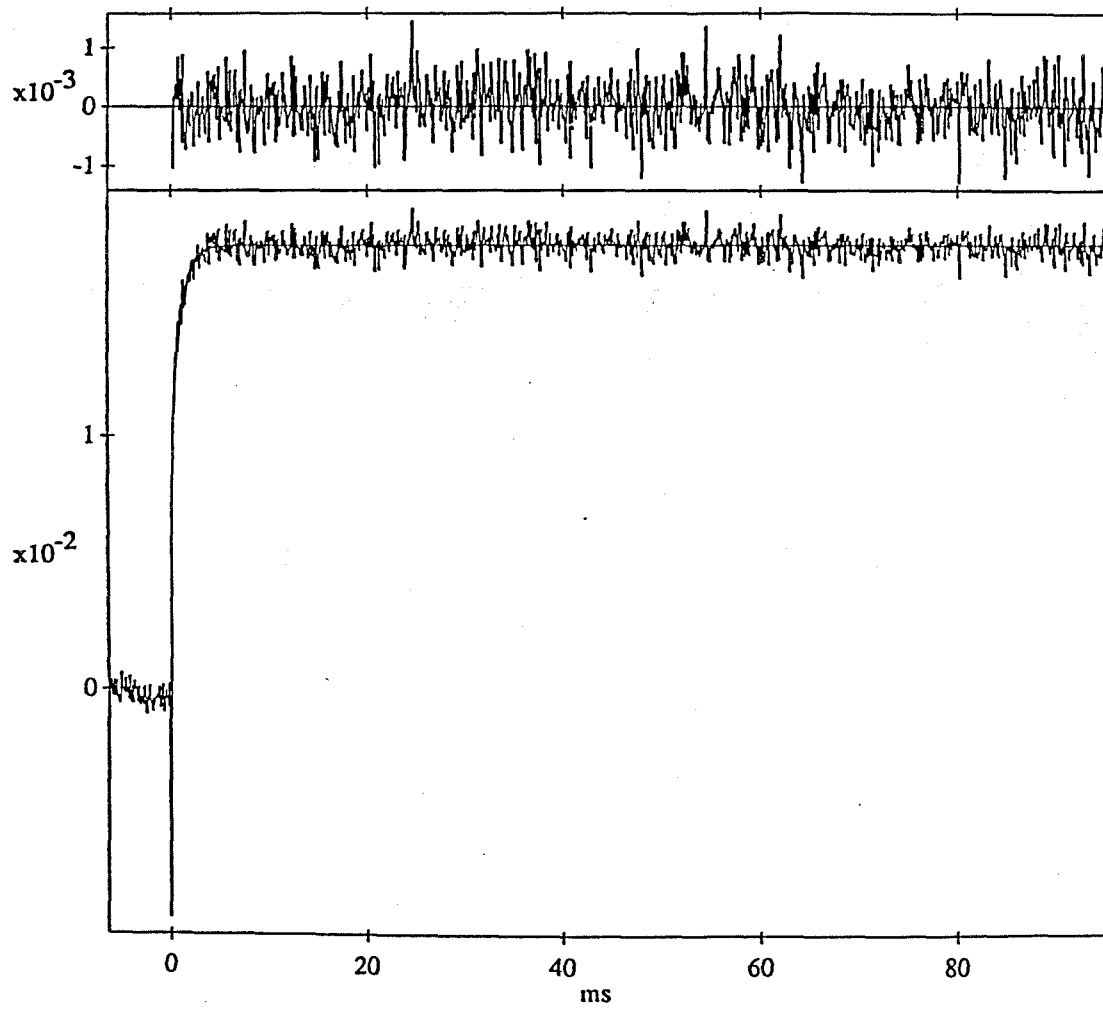
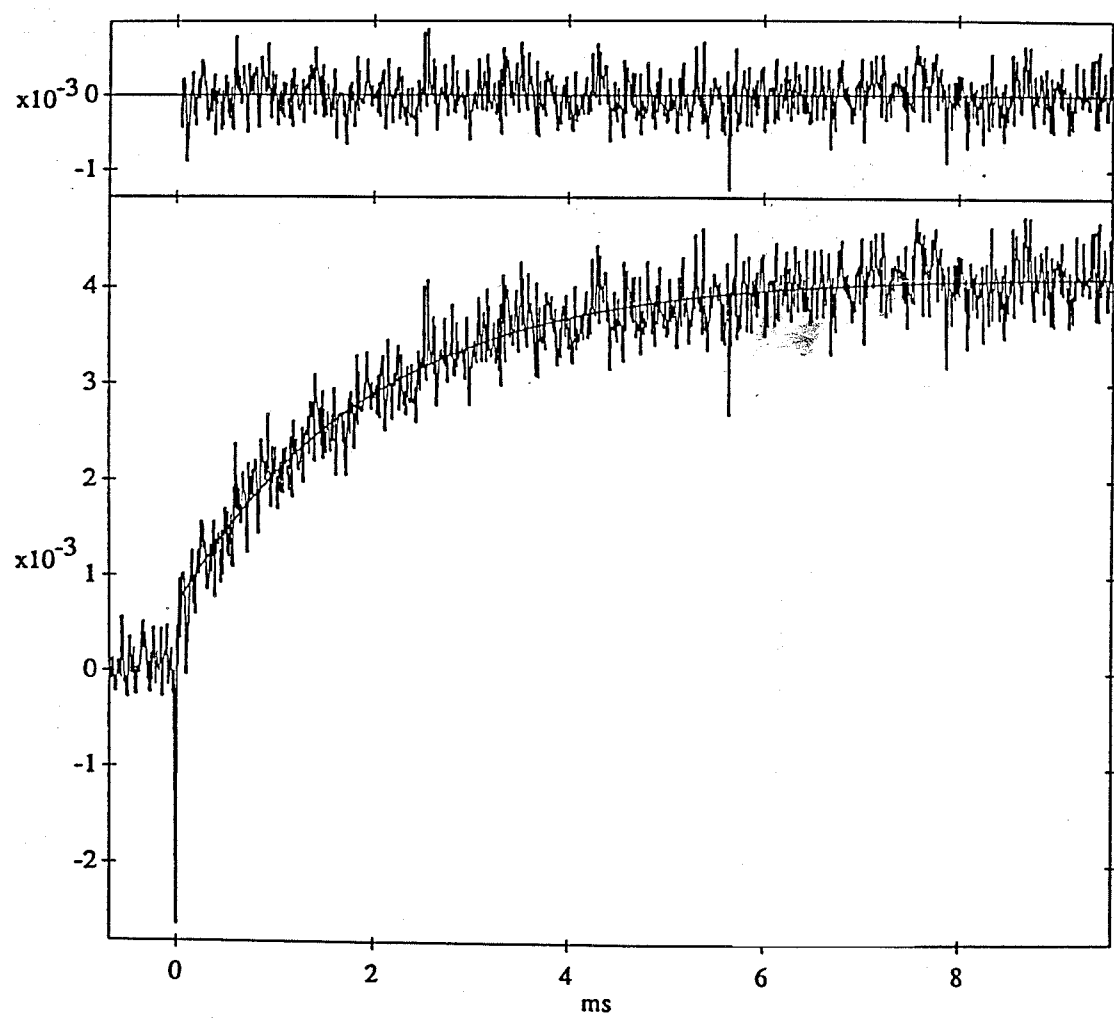
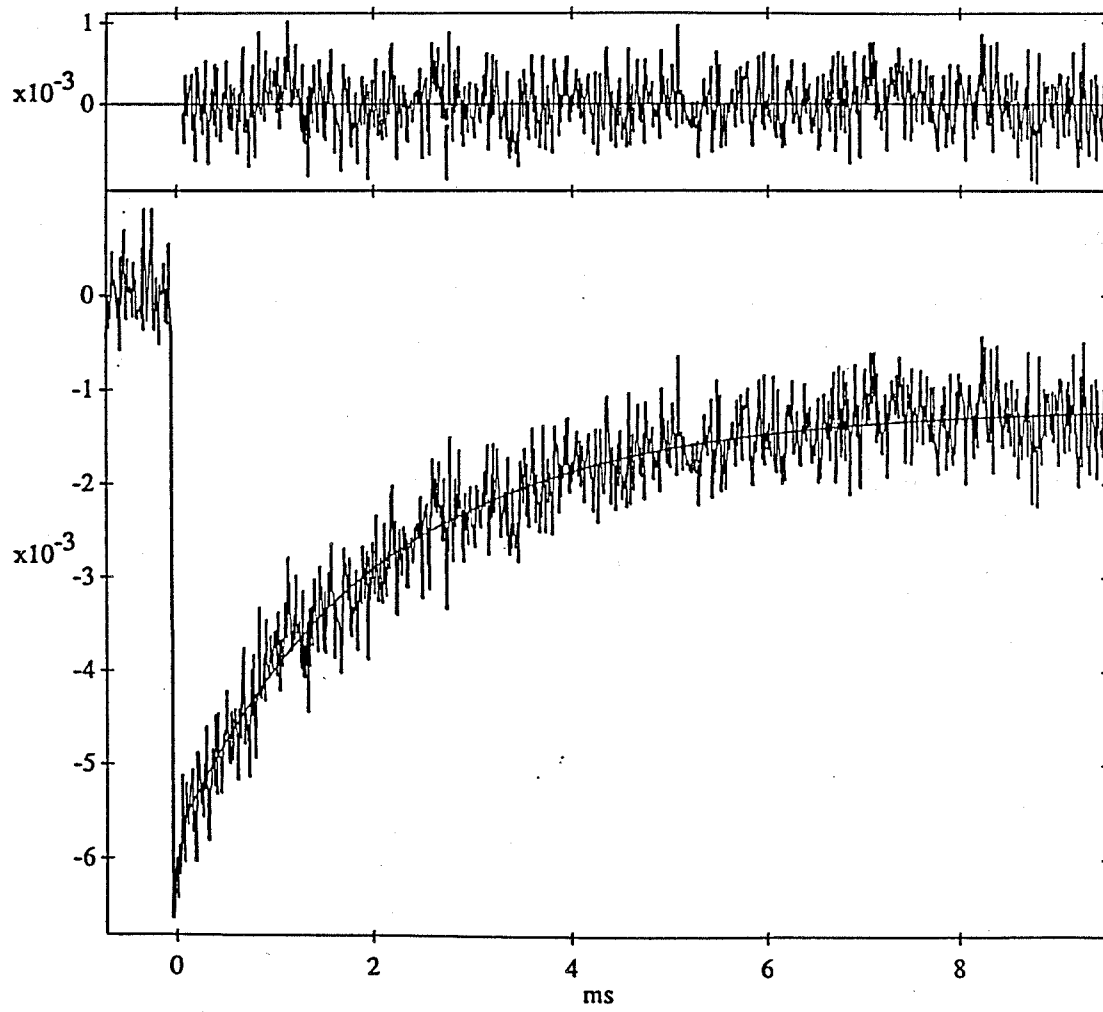


Figure 2.6. Representative kinetics from flash-quench experiments with H126 mutant. $\text{Co}(\text{NH}_3)_6$ was used as quencher in those experiments. Traces are from different concentrations. (a) The buildup of Cu^{2+} is monitored at 625 nm; the concentration is 60 μM . (b) The buildup of Cu^{2+} is monitored at 625 nm; the concentration is 20 μM . (c) The reduction of Ru^{3+} is monitored by the changes in absorption at 500 nm; the concentration is 20 μM . (d) The reduction of Ru^{3+} is monitored by the changes in absorption at 500 nm; the concentration is 5 μM .







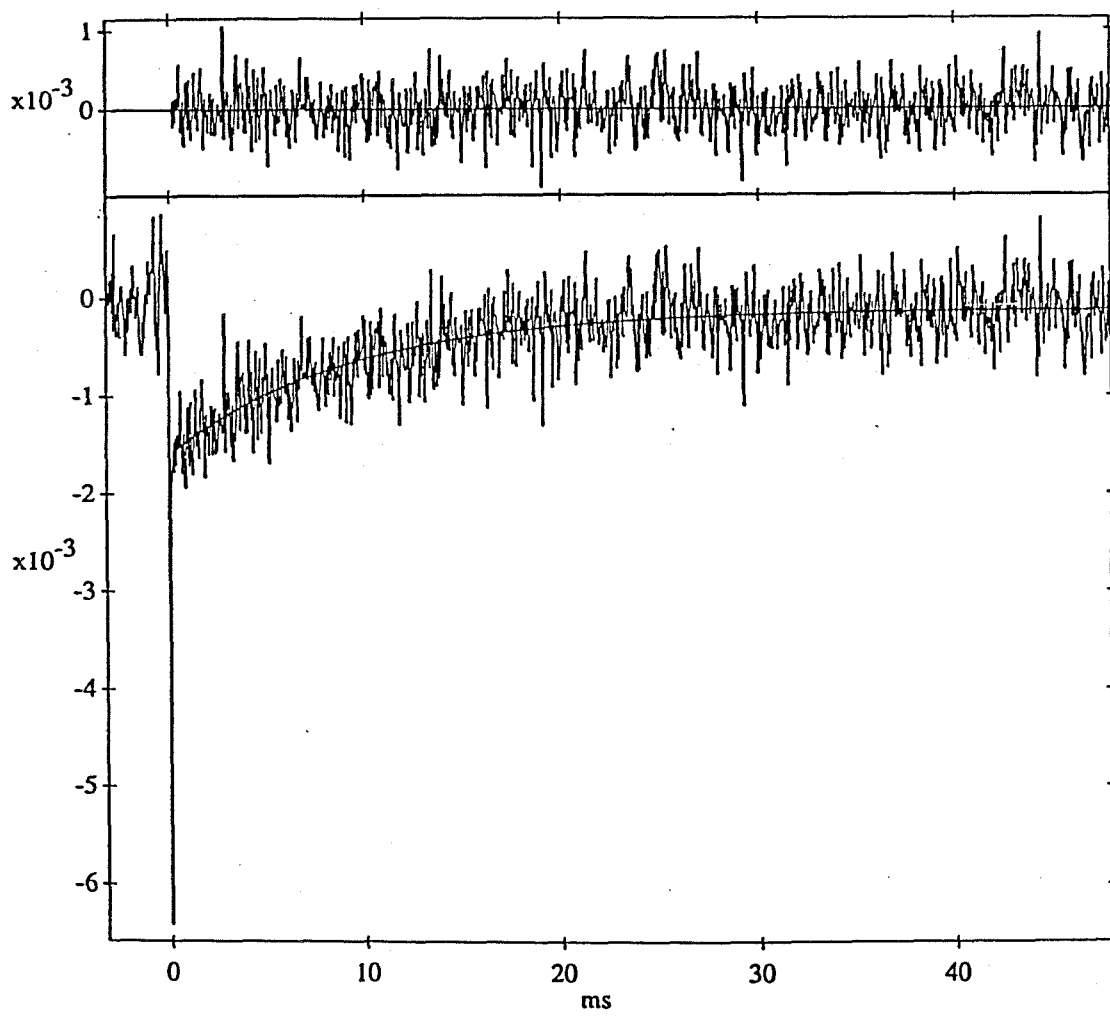


Figure 2.7. Concentration dependence of the ET kinetics for the H126 mutant. A rate constant of 130/s is obtained by extrapolation to infinite dilution. The individual points represent average values for typically 5 to 10 measurements per concentration.

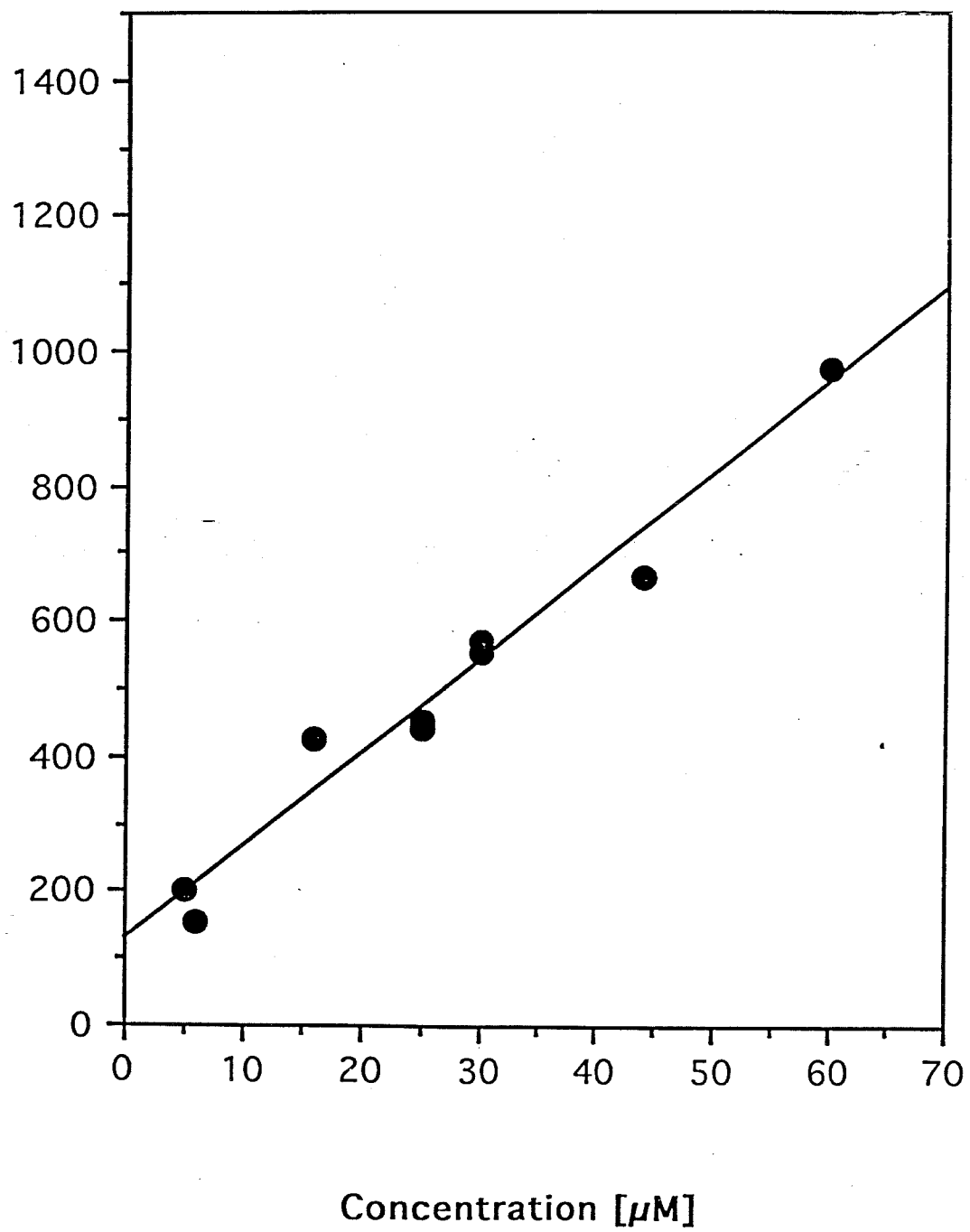


Figure 2.8. Temperature dependence of the ET kinetics for the H124 mutant.

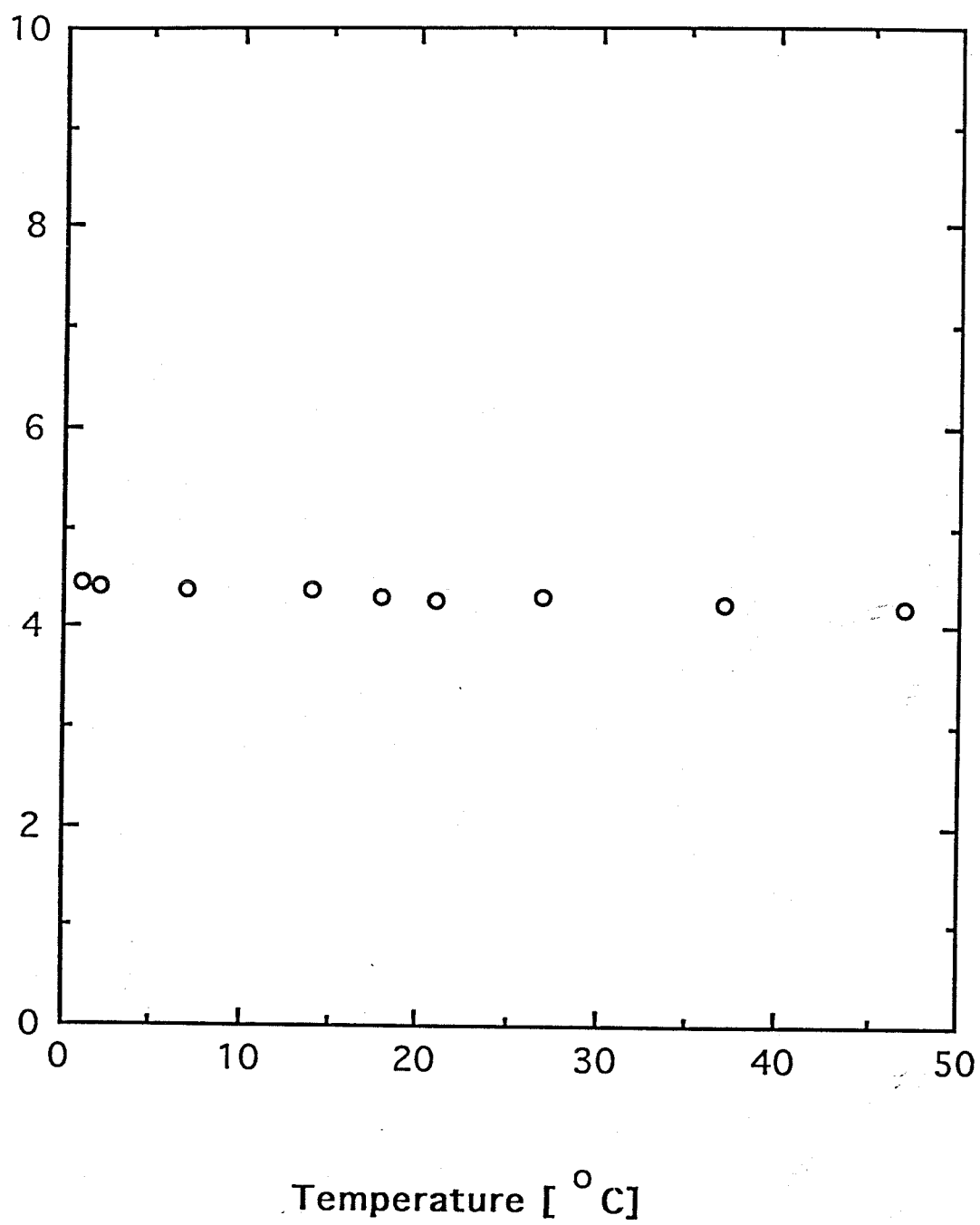


Figure 2.9. Sequence of the ruthenium-containing fragment for the H124 tryptic digest, as determined by peptide sequencing. Plotted are the individual intensities with which the amino acids were identified.

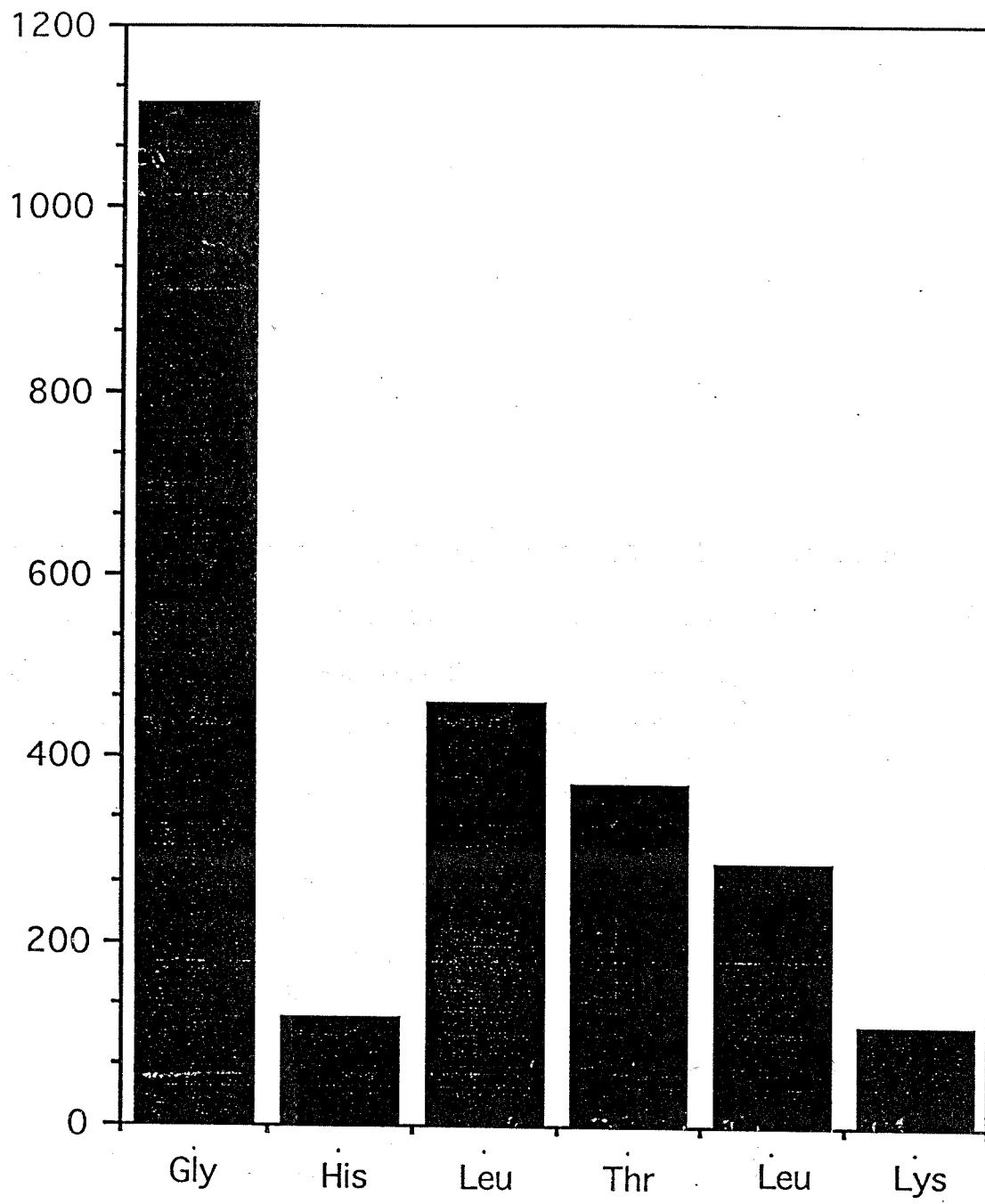


Figure 2.10. Plot of $\log k_{\max}$ vs. metal-metal separation distance (R_M) in Ru-modified azurin. The solid line indicates the best linear fit to the data; the dotted line shows the distance dependence predicted for an idealized β -strand; and the dashed line shows the distance dependence predicted by Dutton and coworkers. R_0 , the close contact distance, is taken to be 3 Å. Inset: observed kinetics of Cu^{2+} formation by $\text{Cu}^+ \rightarrow \text{Ru}^{3+}$ ET.

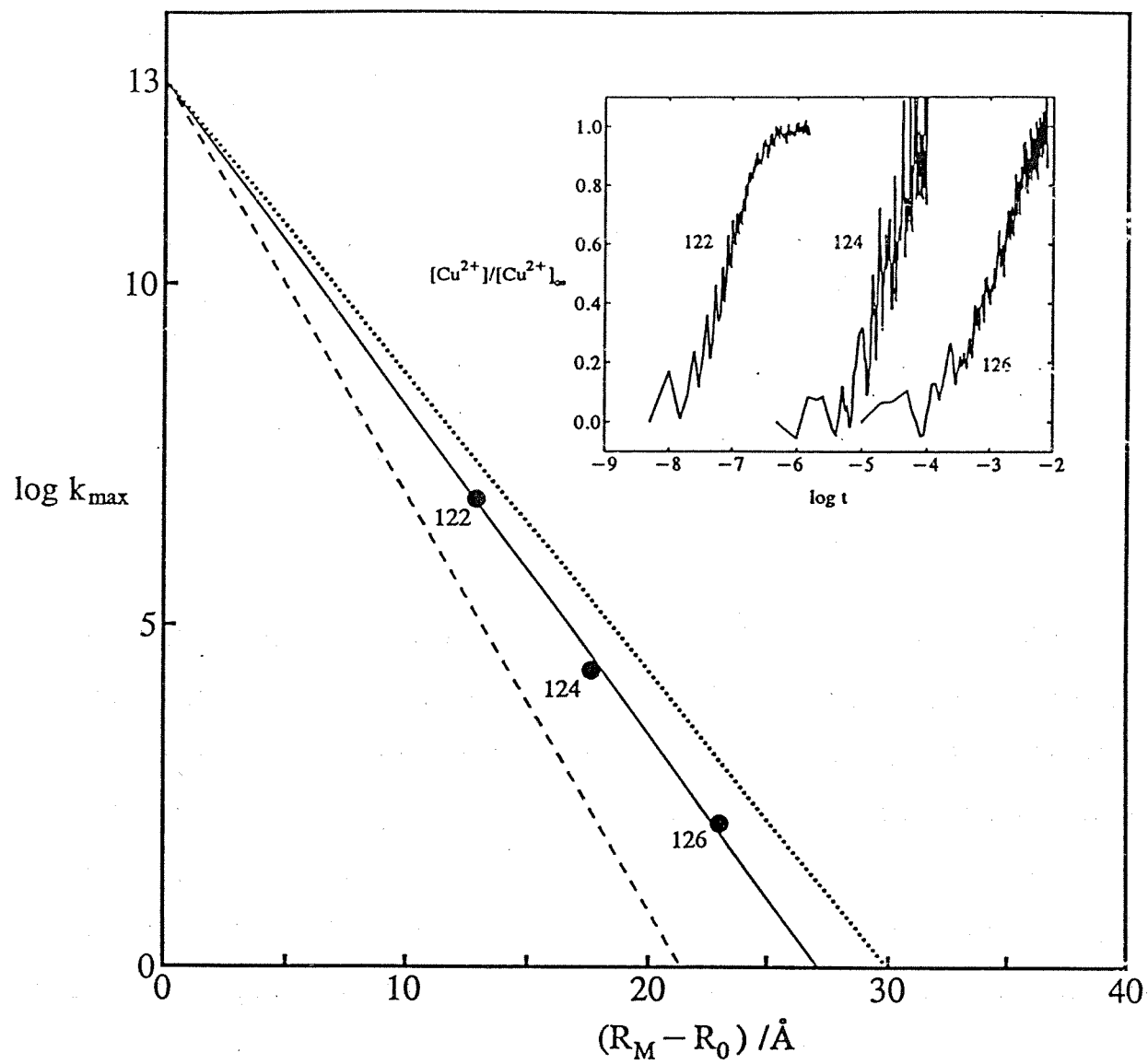
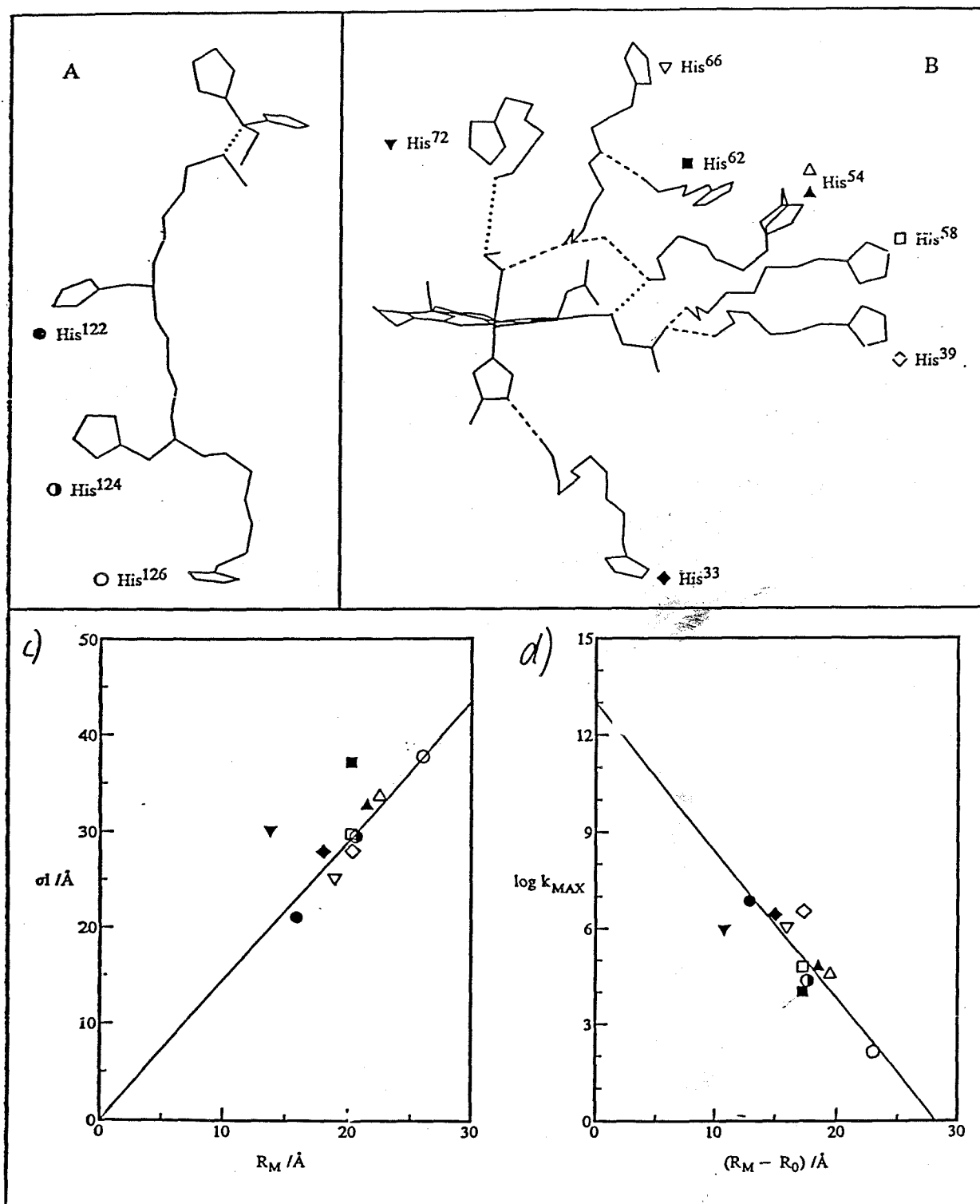


Figure 2.11. **Upper:** tunneling pathways from surface His residues to metal centers in azurin (A) and cytochrom *c* (B). Solid lines indicate covalent links, dashed lines indicate H-bonds, and dotted lines indicate through-space contacts. **Lower left:** plot of σ_s vs. metal-metal separation distances (R_M) in Ru-modified azurin and cytochrome *c*. Solid line is the best linear fit to all but the His⁶² and His⁷² data points, and is constrained to pass through the origin. **Lower right:** plot of $\log k_{\max}$ vs. metal-metal separation distance (R_M) for the two Ru-modified proteins. Solid line is that predicted from the σ_s vs. R_M plot.



Chapter 3

Electronic Coupling through the C112-Strand in Azurin

Abstract

Donor-acceptor electronic coupling through a β -strand with terminal coupling to the copper center via the C112 ligand was investigated by measuring intramolecular electron transfer rates in azurin. Surface histidines were introduced by mutagenesis and modified with an artificial redox site, a $\text{Ru}(\text{bpy})_2\text{im}^{2+}$ -complex. ET from Cu^+ to Ru^{3+} was monitored by time-resolved absorption spectroscopy. The rate constants for the H109 and H107 mutants are $8.0(5) \times 10^5$, and $2.4(3) \times 10^2 \text{s}^{-1}$, respectively. These rates are faster than the respective rates measured for the equidistant modifications on the neighboring M121 strand (Chapter 2). While the M121 ligand is only weakly coordinated to the copper, C112 is a strong ligand. The data presented in this chapter suggest that the stronger copper-ligand interaction in the case of C112 enhances the electronic coupling for electron transfer.

INTRODUCTION

Long range electron transfer through proteins plays an important role in many biological processes. In an extensive study, Dutton and coworkers examined electron transfer (ET) reactions in a variety of proteins [1]. It was found that the ET rates decrease with increasing D-A separation, independent of the exact nature of the intervening protein medium. Dutton suggested that the electronic coupling falls off exponentially with distance, according to Eq.1:

$$H_{AB} = H_{AB}^0 \exp[-\beta(d - d_0) / 2] \quad (1)$$

where H_{AB}^0 is the coupling matrix element at van der Waals contact ($d_0 \sim 3 \text{ \AA}$) and β is the distance decay factor.

Several other theoretical and experimental studies, however, have shown that the nature of the bridge can influence the electronic coupling in biological molecules [2-29]. For example, Wuttke *et al.* observed that ET in cytochrome *c* does not follow a simple distance dependence [24, 25] and can better be described by a model based on specific ET pathways. This model was developed by Beratan and Onuchic [3-5, 28, 29], who used a search algorithm that provides estimates of electronic couplings (ET pathways) in proteins of known structure. In their model, electronic coupling can be mediated very efficiently by covalent bonds, whereas hydrogen bonds and through-space contacts cause a significantly larger decay of the coupling [3-5, 28, 29]. The overall coupling depends on the product of individual decay constants associated with those interactions (ϵ_c , ϵ_h , and ϵ_s respectively). Such pathway calculations also predict that different structural motifs are likely to exhibit significantly different ET properties. For example, properly aligned β -sheets are likely to provide efficient electron tunneling pathways, since they consist of a large number of interconnected covalent tunneling pathways. These predictions were tested in this work that was reported in the previous chapter. It was found that at very

long distances (over 20 Å) coupling is still significant and that the experimentally observed value for β (1.1 Å^{-1}) is close to the predicted 1.0 Å^{-1} .

All of those ET tunneling pathways used the weakly coupled M121 ligand for the terminal coupling to the copper center. It was argued that the weak M121-copper interaction causes a decrease in the observed rate constants. In contrast to M121, C112 is a strong ligand. It is responsible for the blue color of azurins and other blue copper proteins, a color that comes from absorption in the 600 nm region that is attributable to ligand-to-metal charge transfer from the π -orbital on the sulfur to the singly occupied $d_{x^2-y^2}$ on the copper [30-32]. The discovery of this intense charge-transfer band raised the question of the evolutionary importance of the C112 ligand for ET, as it is possible that the strongly coupled C112 ligand provides an efficient tunneling pathway (see, for example, ref. [32] for discussion). Some evidence from experiments involving the related blue copper protein plastocyanin suggests a role of the homologous C82 as a port of entry for electrons into the copper site [33-35]. A similar tunneling pathway for ET between type I and type III Cu-sites has been proposed in ascorbate oxidase[36]. However, the exact nature of ET in blue copper proteins is not fully understood, as several other sites in plastocyanin and azurin appear to play important roles as well [21, 37, 38]. For example, there is evidence that the self-exchange reaction in azurin does not proceed through C112, but rather uses the H117 ligand[37].

In this chapter we address the question of electron tunneling through the C112 ligand. The H109 and H107 positions are next to the H124 and H126 positions on the neighboring M121 strand (Fig. 3.1), which facilitates the comparison of the H109 and H107 rates with the H124 and H126 rates reported in Chapter 2. The rates measured at analogous points on the C112 strand are higher consistent with stronger coupling of the strand to the C112-Cu unit.

MATERIALS AND METHODS

The methodology used in this chapter is very similar to the methodology in Chapter 2 and therefore only summarized here briefly. The proteins were expressed in *E.coli* using a T7 polymerase expression system [39] and purified by FPLC according to previously published protocols [40]. The purified proteins were then incubated with $\text{Ru}(\text{bpy})_2(\text{CO}_3)$ [41]. After ruthenation the proteins were again purified by FPLC. The addition of excess imidazole yielded $\text{Ru}(\text{bpy})_2(\text{im})(\text{His})^{2+}$. Rates of $\text{Cu}^+ \rightarrow \text{Ru}^{3+}$ ET were measured using a previously described flash-quench technique [41]. $^*\text{Ru}(\text{bpy})_2(\text{im})(\text{His})^{2+}$ -azurin was generated using a 480 nm laser pulse (25 ns pulse width). The excited state was quenched by $\text{Ru}(\text{NH}_3)_6^{3+}$ (5 mM solution) in the case of His 109, and by $\text{Co}(\text{NH}_3)_5\text{Cl}^{2+}$ (6 mM solution) in the case of His 107, to yield $\text{Ru}(\text{bpy})_2(\text{im})(\text{His})^{3+}$ -azurin. ET was monitored at 625 ($\text{Cu}^+/\text{Cu}^{2+}$), 500 ($\text{Ru}^{3+}/\text{Ru}^{2+}$), 432 ($\text{Ru}^{3+}/\text{Ru}^{2+}$), and 310 nm ($\text{Ru}^{3+}/\text{Ru}^{2+}$). Since a bimolecular ET component was observed for His 107, its intramolecular ET rate was obtained by extrapolation to infinite dilution based on concentration dependence measurements. All distances were obtained by molecular modeling using the program ENZYME [42]. The final distance was obtained by averaging the distances in snapshots taken at 1 ps intervals during 200 ps of a constrained dynamics simulation. In all cases, convergence was reached after approximately 20-30 ps of simulation.

RESULTS

DNA

As in Chapter 2 the mutagenesis reactions were carried out with the Muta-Gene kit purchased from Bio-Rad [43, 44]. Once again, the DNA was constructed by simultaneous use of two mutagenic oligonucleotides. The yields for the successful incorporation of both mutations for the mutagenesis was 20 to 30 %.

Protein Expression and Purification

The level of protein expression was comparable to those levels observed in the previous chapter. The yields ranged from 30 to 50 mg/L of bacterial culture for both mutants. Again, the rate of copper uptake was significantly accelerated by heating the protein in 25 mM NaOAc (pH4.5) to 45 °C in experiments in which CuSO₄ is 5 to 10 mM.

Ruthenium Modification

The Ru(bpy)₂(H₂O)(His)-azurin derivatives were purified with a MonoS column. The yield for the H107 mutant was approximately 30%. The H109 mutant, however, had a somewhat lower yield (4 to 5 %). The low yield for the H109 mutant could be related to the sterically less accessible environment around the H109.

Rate Measurements

The kinetics traces for the H107 and the H109 mutants are shown in Fig. 3.2 and Fig. 3.1. The rates for the H109 mutant were measured independently by photoinduced and flash-quench techniques. Both techniques resulted in the same rate constant for Cu⁺ to Ru³⁺ electron transfer, $8.0(5) \times 10^5 \text{ s}^{-1}$. For the H107 mutant, no significant charge separation was observed and the signal sizes were too small to measure the intramolecular ET rate with the photoinduced technique. For this mutant the rate constant was obtained only by the flash-quench method. Unlike the case of the H109 mutant, the recombination rate between Ru(NH₃)₆²⁺ and Ru³⁺ was in the same range as the rate of the intramolecular

ET reaction. However, those data could be fitted to a model that considers both processes simultaneously and the ET rates could be extracted. Consistent results were obtained using the irreversible quencher $\text{Co}(\text{NH}_3)_5\text{Cl}^{2+}$. As was the case for the H126 mutant, the intramolecular rate constant had to be extracted by extrapolation to infinite dilution as the fitted ET rate constants were concentration dependent. The extracted rate constant is $2.4(3) \times 10^2 \text{ s}^{-1}$ (Fig. 3.4).

DISCUSSION

In this chapter we investigate the electron tunneling through a β -strand that runs parallel to the M121 strand discussed in Chapter 2 (Fig. 3.1). As pointed out in that chapter the driving force for the ET reactions considered here is nearly equal to the estimated reorganization energy, and the observed rates are expected to be within 5 % of the activationless rates (k_{max}).

A plot of $\log(k_{\text{max}})$ vs. R_m (metal-metal separation) is shown in Fig.3.5. As in the previous chapter, it is nearly linear for all azurin mutants. The best fit line when compared to Chapter 2 is hardly affected by the addition of the two new points. The slope (β) changes from 1.10 to 1.09 \AA^{-1} , which is in even closer agreement with the predicted decay constant for a β -strand (1.0 \AA^{-1}). This analysis corroborates our previous conclusion (Chapter 2) that a β -strand represents a very efficient bridge for electron tunneling.

The strong coupling through the β -strands of azurin can be understood in terms of a Beratan-Onuchic type pathway analysis. All mutants exhibit dominant electron tunneling pathways consisting of covalent bonds. In the case of the backbone (a central part of these pathways) the bonds are aligned such that they provide a straight covalent connection, thereby providing a very efficient bridge for electron tunneling. For covalent interactions,

per-bond decay constants of $\epsilon_c=0.6$ are used in the Beratan-Onuchic pathway model [3-5, 28, 29]. This normal per-bond decay constant seems appropriate for all bonds except the ligand-copper interactions, as the M121-Cu coupling is expected to be weaker than the C112-Cu coupling. The M121 ligand is only weakly coordinated (bond length ~ 3.2 Å)[20-22, 45], and it can easily be replaced by mutagenesis without significantly altering the characteristics of the copper site [40, 46]. In contrast, C112 is a strong ligand and mutagenesis at this position completely changes the spectroscopic characteristics of the copper-binding site [47]. A plot of the rate constants (k_{max}) versus tunneling length (σl) is shown in Fig. 3.6. Rather than a physical distance, σl is a measure of the number of carbon-carbon bond equivalents in the ET pathway. Since the per-bond decay in this model is calibrated to be 0.6, it follows that the slope has to be $0.73/\text{Å}$. Without including the ligand-copper interaction into the overall tunneling length, the points are low and would give an intercept of less than 12^\dagger . If we assume the tunneling length for the C112-copper interaction to be the same as for all other covalent bonds treated in this analysis (1.4 Å), we can fit the two resulting points from the C112 strand to a best fit line with a required slope of $0.73/\text{Å}$. If we now treat the M121-copper interaction as an adjustable parameter in order to get the best agreement to the best fit line, we obtain a tunneling length of 3.9 Å.

How reasonable are those two numbers? Using self-consistent-field-X α -scattered-wave calculations, the strength of the Cu-M121 interaction has been estimated to be 30% of a "normal covalent bond"[31]. If we assume that the decay constant for this bond is 30% of that of a normal covalent bond ($\epsilon_{Cu-S}=0.18$), the copper-Met interaction is equivalent to 3.35 covalent carbon-carbon bonds, yielding a σl of 4.69 Å. In the Beratan-Onuchic pathway model, we can correct for the long bond length by defining the interaction as a

[†] This would correspond to a rate constant of (probably significantly) less than 10^{12}s^{-1} at the adiabatic limit which should be reached at values of σl larger than 0 Å. From theoretical consideration a rate of 10^{13}s^{-1} would be expected.

through space contact. In this representation, the coupling falls off exponentially as the copper-sulfur distance increases from a normal copper-sulfur bond (~ 2.3 Å) according to Eq. 2:

$$\epsilon_S = 0.6 \exp[-1.7 (R_{\text{Cu-S}} - 2.3 \text{ Å})] \quad (2)$$

where $R_{\text{Cu-S}}$ is the copper-sulfur bond length in Å. Using this approach, we find that the Met121-copper interaction exhibits an electronic coupling equivalent to 4 covalent carbon-carbon bonds (i.e., an effective $\sigma l = 4 \times 1.4 \text{ Å} = 5.6 \text{ Å}$). The estimate for the M121-copper interaction derived from Fig. 3.6 is in very close agreement with the results from the self-consistent-field- $X\alpha$ -scattered-wave calculations, while it is somewhat smaller than what one would expect from a through space interaction.

The difference in σl (2.5 Å) for the C112-copper and M121-copper bond roughly corresponds to a 6-fold rate enhancement for the C112-copper interaction. This rate enhancement might be somewhat smaller than expected. It is important to note that the charge transfer from C112 to copper has been assigned as a $\text{SCys}\pi\text{-Cu } d_{x^2-y^2}$ transition. However, according to our model, tunneling occurs mainly through σ -bonds, and the σ - π coupling at the sulfur atom might be relatively weak diminishing what could have been a much stronger rate enhancement.

So far the discussion has mainly focused on the single pathway picture. According to a recent theoretical analysis by Regan *et al.* [48], this is sufficient for H122 and H124. For those mutants only, the pathway through the M121 ligand appears to be dominant. However, in the case of H126, a secondary route through C112 emerged from the calculations. The Regan study [48] did not analyze the tunneling pathways for H109 and H107. However, it appears that the single pathway model should be valid in those cases. That is, as found for the M121 strand, straight covalent ET pathways should dominate. Additionally, in the case of H107 and H109, the covalent bonds are linked to the

strongest ligand, C112, and the resulting tunneling pathways should therefore be better than all others. Such a description is fully consistent with the data obtained in this chapter, as can be seen from the rate ratios for the respective (almost equidistant) mutants on the Met and the Cys strands. Although the rates for the Cys strand are always greater than those for the corresponding positions on the Met strand, the H109/H124 ratio is ~ 30 , whereas the H107/H126 ratio is only ~ 2 . The latter ratio appears to be lowered because the H126 rate is higher than predicted by a single pathway model. The availability of constructively interfering pathways from Cu to H126 is the likely explanation of the enhanced coupling [48].

As shown in the previous chapter, the tunneling-pathway model makes it possible to compare ET rate constants measured for azurin to results obtained for other Ru-modified proteins. Fig. 3.7 shows that the addition of the H107 and the H109 rates (and the somewhat shortened tunneling length for the Cu-Met interaction) to a plot of R_m vs. σ does not affect the overall correlation with the best fit line from Chapter 2. Again, only 2 of the cytochrome *c* rates, H62 and H72, exhibit larger deviations from the best fit line*. One would therefore predict only those two rates to deviate from the predicted best fit line ($\beta=1.06/\text{\AA}$) in a plot of $\log(k_{max})$ vs. R_m . This prediction is reasonably consistent with the data (Fig. 3.8). The best fit (excluding H62 and H72) would have a slope of $1.07/\text{\AA}$, which is very close to the predicted value.

In summary, it appears that electron tunneling through C112 is more efficient than through the M121 ligand. However, the rate enhancement due to the stronger bond seems at most to be an order of magnitude.

* The best fit line using the data from this chapter does not change the slope significantly. It changes from 1.45 (chapter 2) to 1.39. This can be converted to a predicted $\log(k_{max})$ vs. R_m decay constant β of $1.0/\text{\AA}$ rather than $1.06/\text{\AA}$.

BIBLIOGRAPHY

1. Moser, C.C., Keske, J.M., Warncke, K., Farid, R.S., and Dutton, P.L., *Nature* **355**, 796 (1992)
2. Friesner, R.A., *Structure* **2**, 339 (1994)
3. Onuchic, J.N., and Beratan, D.N., *J. Chem. Phys.* **92**, 722 (1990)
4. Onuchic, J.N., Andrade, P.C.P., and Beratan, D.N., *J. Chem. Phys.* **95**, 1131 (1991)
5. Onuchic, J.N., Beratan, D.N., Winkler, J.R., and Gray, H.B., *Annu. Rev. Biophys. Biomol. Struct.* **21**, 349 (1992)
6. Skourtis, S.S., Regan, J.J., and Onuchic, J.N., *J. Phys. Chem.* **98**, 3379 (1994)
7. Siddarth, P., and Marcus, R.A., *J. Chem. Phys.* **94**, 2985 (1990)
8. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **94**, 8430 (1990)
9. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **96**, 3213 (1992)
10. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **97**, 2400 (1993)
11. Siddarth, P., and Marcus, R.A., *J. Phys. Chem.* **97**, 13078 (1993)
12. Stuchebrukhov, A.A., and Marcus, R.A., *J. Phys. Chem.* in press
13. Gruschus, J.M., and Kuki, A., *J. Phys. Chem.* **97**, 5581 (1993)
14. Broo, A., and Larsson, S., *J. Phys. Chem* **97**, 5581 (1993)

15. Christensen, H.E.M., Conrad, L.S., Mikkelsen, K.V., and Ulstrup, J., *J. Phys. Chem* **96**, 4451 (1992)
16. Evenson, J.M., and Karplus, M., *Science* **262**, 1247 (1993)
17. Farver, O., and Pecht, I., *J. Am. Chem. Soc.* **114**, 5764 (1992)
18. Farver, O., Skov, L.K., Pascher, T., Karlsson, G., Nordling, M., Lundberg, L.G., Vänngård, T., and Pecht, I., *Biochemistry* **32**, 7317 (1993)
19. Canters, G.W., and van de Kamp, M., *Curr. Opin. Struct. Biol.* **2**, 859 (1992)
20. Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G.W.J., *J. Mol. Biol.* **221**, 765 (1991)
21. Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G.W.J., *J. Mol. Biol.* **218**, 427 (1991)
22. Nar, H., Huber, R., Messerschmidt, A., Filippou, A.C., Barth, M., Jaquinod, M., van de Kamp, M., and Canters, G.W., *Eur. J. Biochem* **205**, 1123 (1992)
23. Karpishin, T.B., Grinstaff, M.W., Komar-Panicucci, S., McLendon, G., and Gray, H.B., *Structure* **2**, 415 (1994)
24. Wuttke, D.S., Bjerrum, M.J., Winkler, J., and Gray, H.B., *Science* **256**, 1007 (1992)
25. Wuttke, D.S., Bjerrum, M.J., Chang, I-Jy, Winkler, J.R., and Gray, H.B., *Biochim. Biophys. Acta* **1101**, 168 (1992)
26. Casimiro, D.R., Wong, L.-L., Colon, J.L., Zewert, T.E., Richards, J.H., Chang, I-J., Winkler, J.R., and Gray, H.B., *J. Am. Chem. Soc.* **115**, 1485 (1993)
27. Casimiro, D.R., Richards, J.R., Winkler, J.R., and Gray, H.B., *J. Phys. Chem* **97**, 13073 (1993)

28. Beratan, D.N., Betts, J.N., and Onuchic, J.N., *Science* **252**, 1285 (1991)
29. Beratan, D.N., Betts, J.N., and Onuchic, J.N., *J. Phys. Chem.* **96**, 2852 (1992)
30. Solomon, E.I., Hare, J.W., Dooley, D.M., Dawson, J.H., and Stephens, P.J., *J. Am. Chem. Soc.* **102**, 168 (1980)
31. Lowery, M.D., and Solomon, E.I., *Inorg. Chim. Acta* **200**, 233 (1992)
32. Solomon, E.I., and Lowery, M.D., *Science* **259**, 1575 (1993)
33. He, S., Modi, S., and Gray, J.C., *EMBO* **10**, 4011 (1991)
34. Sykes, A.G., *Chem. Soc. Rev.* **14**, 283 (1985)
35. Sykes, A.G., *Struct. Bonding* **75**, 175 (1990)
36. Messerschmidt, A., Ladenstein, R., Huber, R., Bolognesi, M., Avigliano, L., Petruzzelli, R., Rossi, A., and Finazzi Agro, A., *J. Mol. Biol.* **224**, 179 (1992)
37. van de Kamp, M., Silvestrini, M.C., Brunori, M., van Beeumen, J., Hall, F.C., and Canters, G.W., *Eur. J. Biochem.* **194**, 109 (1990)
38. Nordling, M., Sigfridsson, K., Young, S., Lundberg, L.G., and Hansson, O., *FEBS Lett* **291**, 327 (1991)
39. Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorf, J.W.: *Methods in Enzymology*, San Diego: Academic Press; 1990.
40. Chang, T.K., Iverson, S.A., Rodrigues, C.G., Kiser, C.N., Lew, A.Y., Germanas, J.P., and Richards, J.H., *Proc. Natl. Acad. Sci.* **88**, 1325 (1991)
41. Chang, I.-J., Gray, H.B., and Winkler, J.R., *J. Am. Chem. Soc.* **113**, 7056 (1991)
42. Lee, F.S., Chu, Z.T., and Warshel, A., *J. Comp. Chem.* **14**, 161 (1993)

43. Kunkel, T.A., *Proc. Natl. Acad. Sci.* **82**, 488 (1985)
44. Kunkel, T.A., Bebenek, K., and McClary, J., *Methods Enzymol.* **204**, 125 (1991)
45. Adman, E.T., Jensen, L.H., *Isr. J. Chem.* **21**, 8 (1981)
46. Karlsson, B.G., Nordling, M., Pascher, T., Tsai, L.-C., Sjölin, L., and Lundberg, L.G., *Protein Engineering* **4**, 343 (1991)
47. Mizoguchi, T.J., DiBilio, A.J., Gray, H., B., and Richards, J.H., *J. Am. Chem. Soc.* **114**, 10076 (1992)
48. Regan, J.J., DiBilio, A.J., Langen, R., Skov, L.K., Winkler, J.R., Gray, H.B., and Onuchic, J.N., *submitted to Angewandte Chemie*

Figure 3.1. Mutants, whose electron transfer pathways are coupled to the copper center via M121 on the right half (referred to as the M121 strand), and mutants, whose electron transfer pathways are coupled to the copper center via C112 on the left half (referred to as the C112 strand).

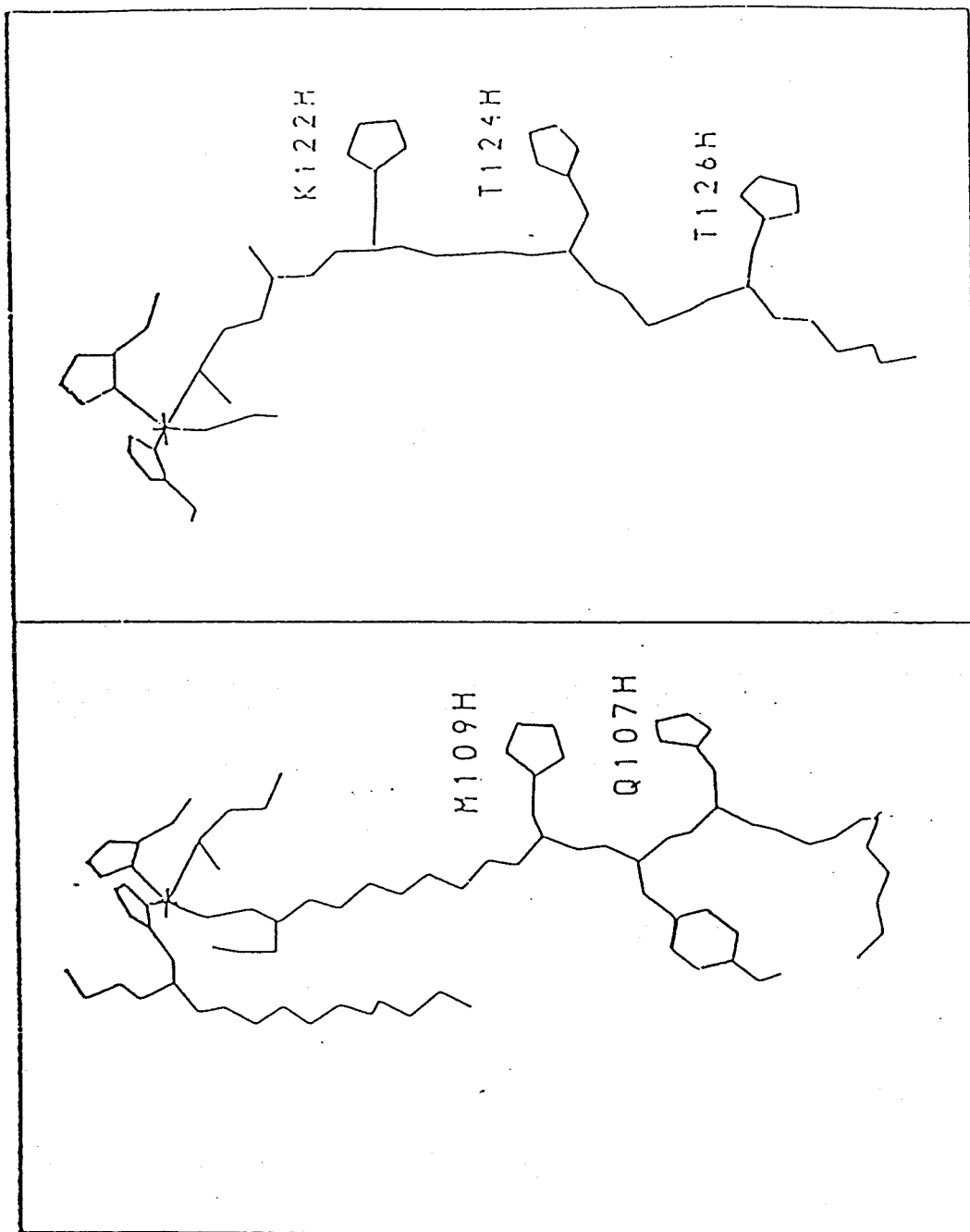
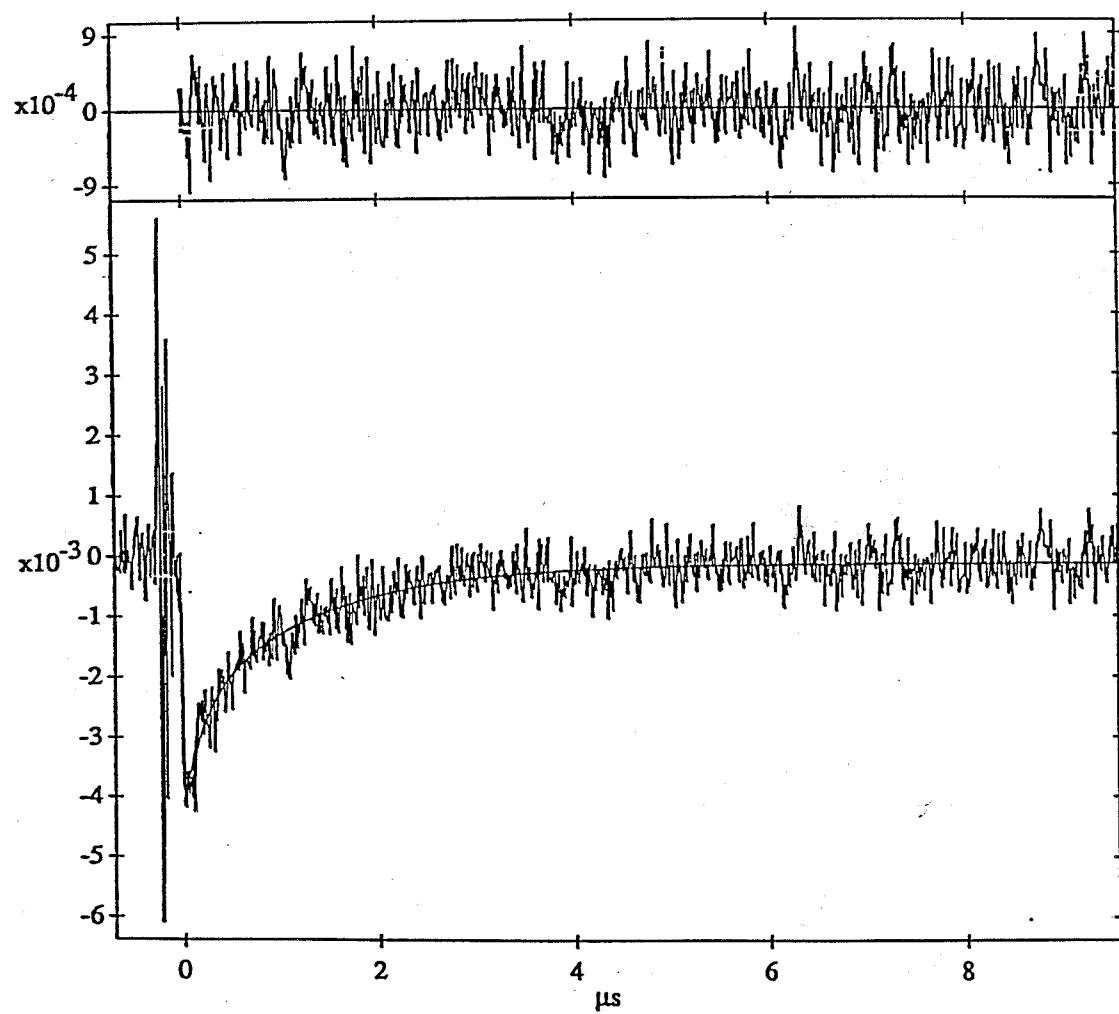
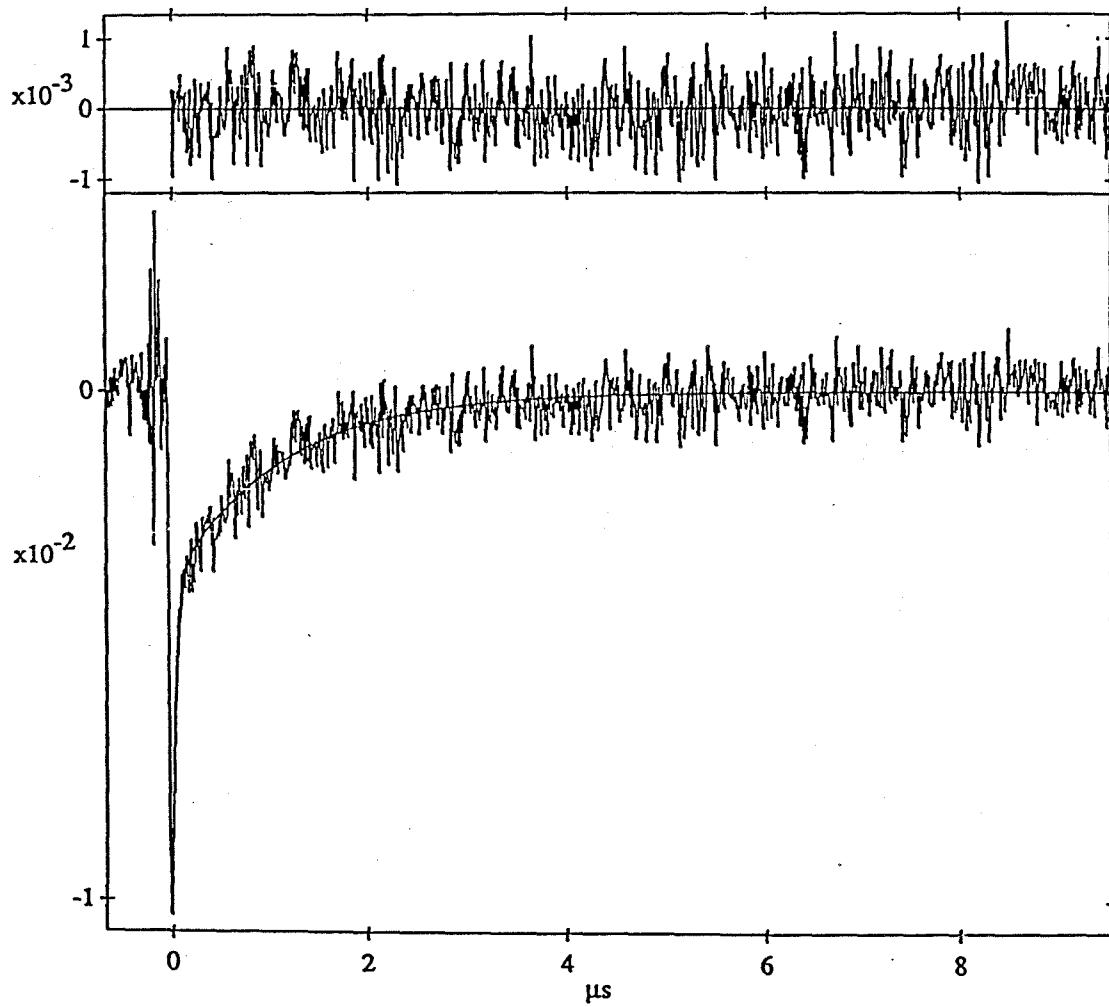
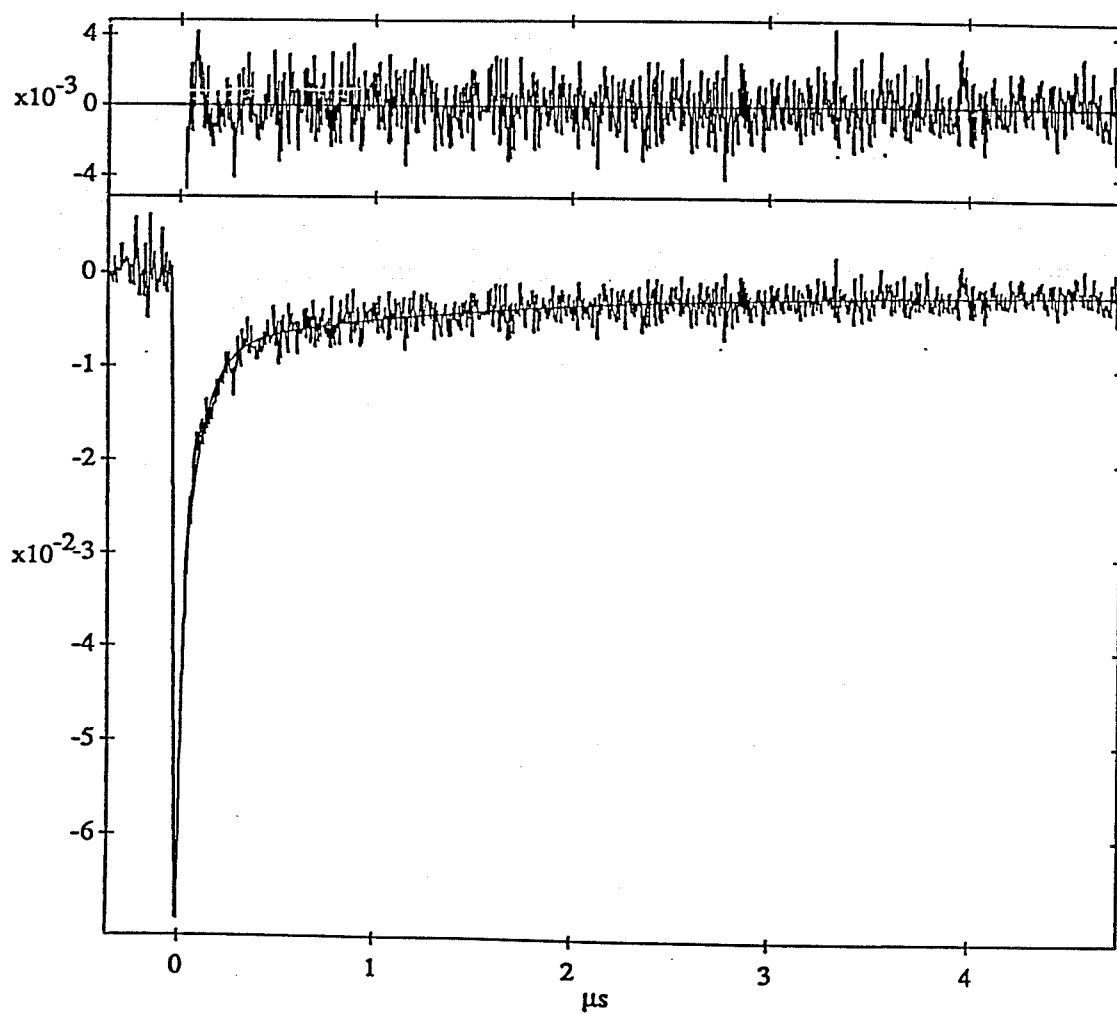


Figure 3.2. Representative kinetics from photoinduced (a, b) and flash-quench (c, d) experiments with the H109 mutant. (a) Photoinduced experiment monitored by the changes in absorption at 432 nm. The bleach and the subsequent recovery is due to the formation of Ru^{3+} followed by the back electron transfer yielding Ru^{2+} (b) Photoinduced experiment monitored by the changes in absorption at 625 nm. The bleach and the subsequent recovery is due to the formation of Cu^{1+} followed by the back electron transfer yielding Cu^{2+} . The initial fast component is due to excited state emission. (c) Flash-quench experiment monitored by the changes in absorption at 500 nm. The bleach and the subsequent recovery are due to the formation of Ru^{3+} followed by electron transfer yielding Ru^{2+} . The initial fast component is due to excited state. (d) Flash-quench experiment monitored by the changes in absorption at 625 nm. The initial fast component (bleach) is due to excited state emission. The subsequent increase in absorbance is due to the formation of Cu^{2+} . All data gave a rate constant for the $\text{Cu}^{+} \rightarrow \text{Ru}^{3+}$ ET reaction of $8 \times 10^5 \text{ s}^{-1}$.







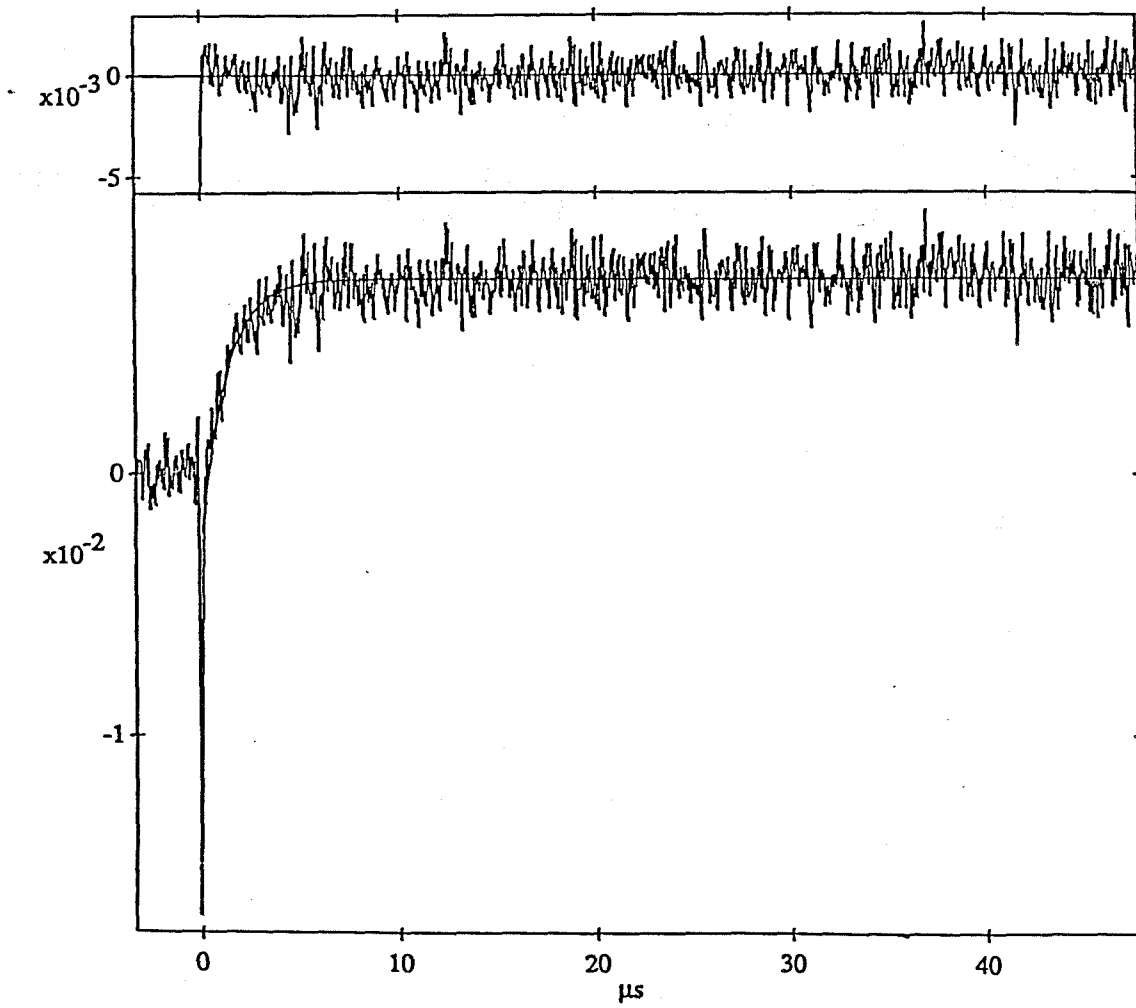
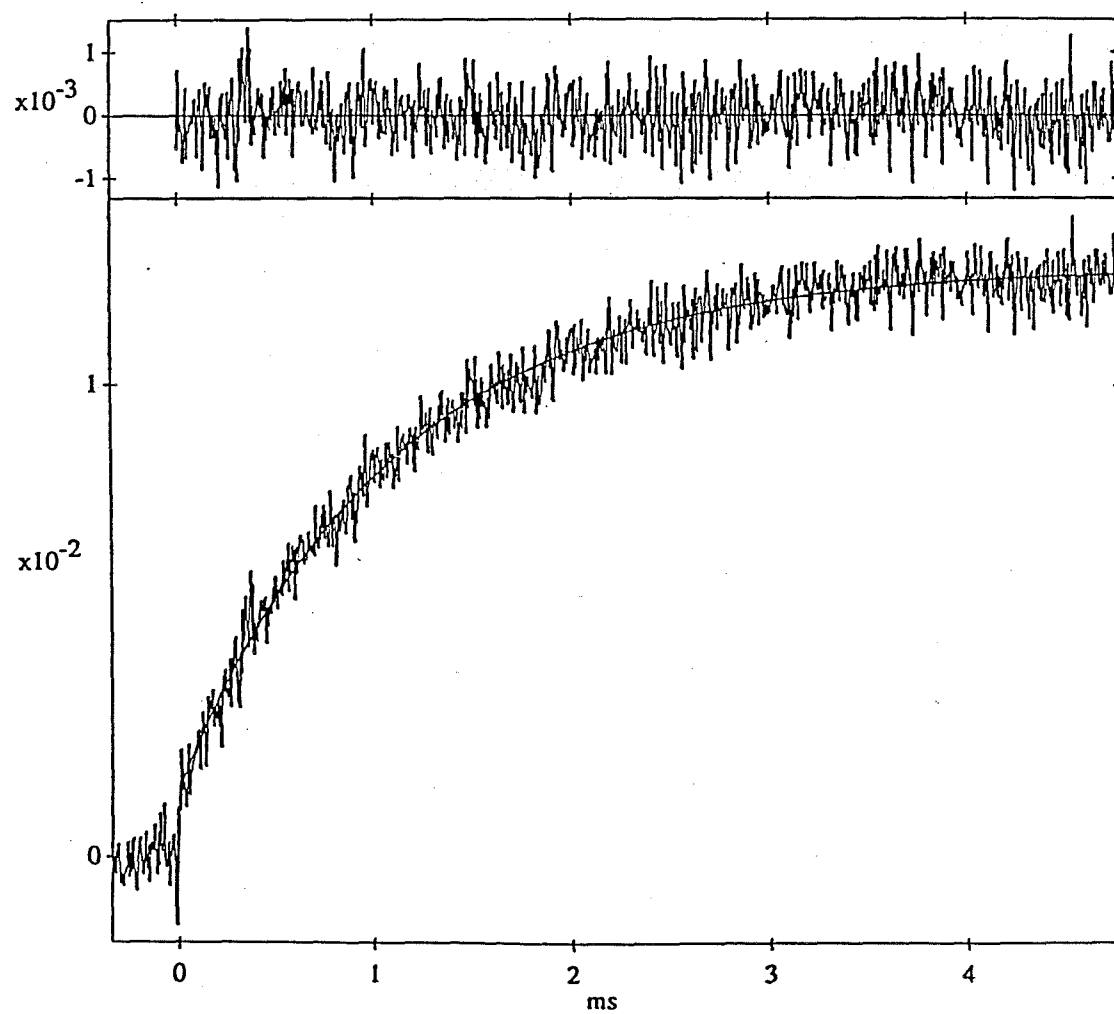
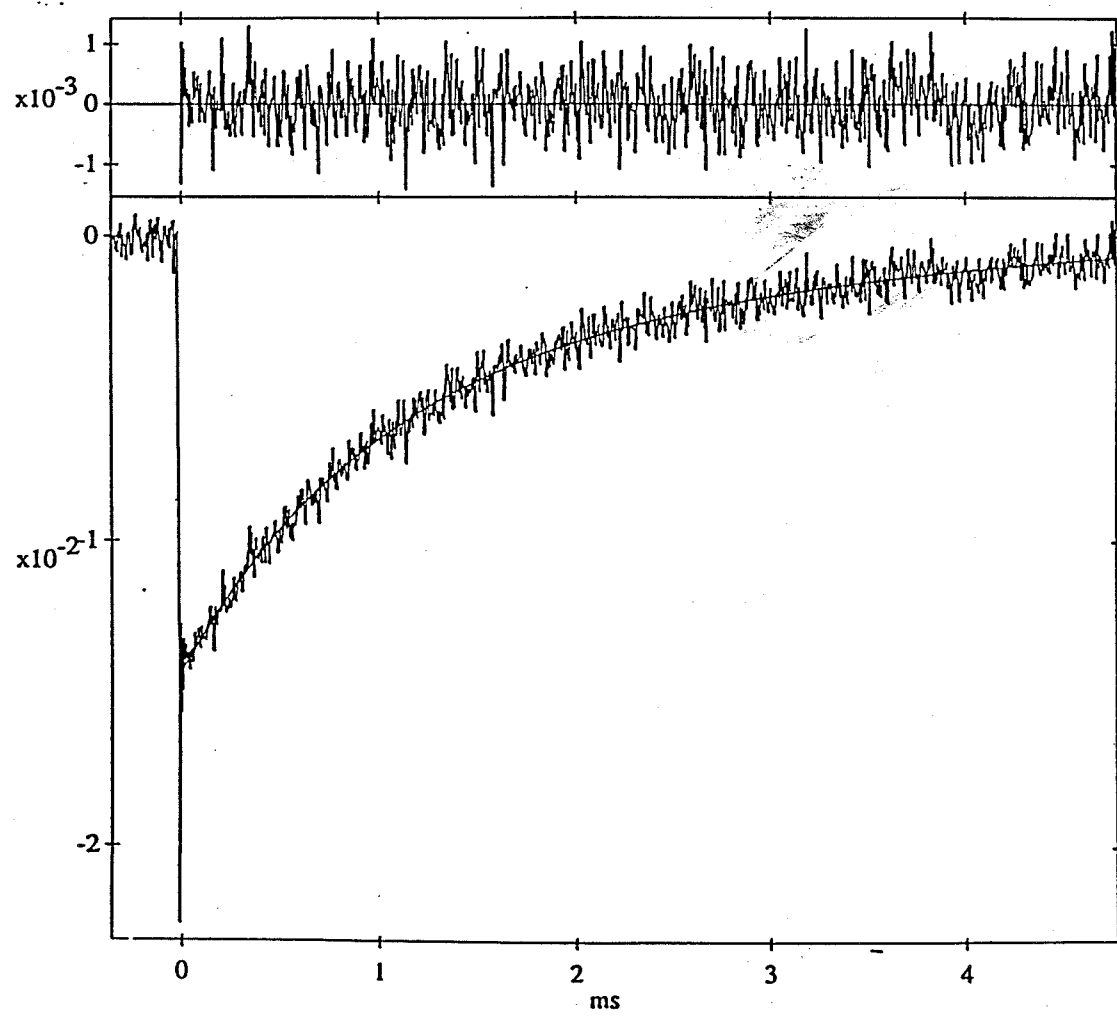
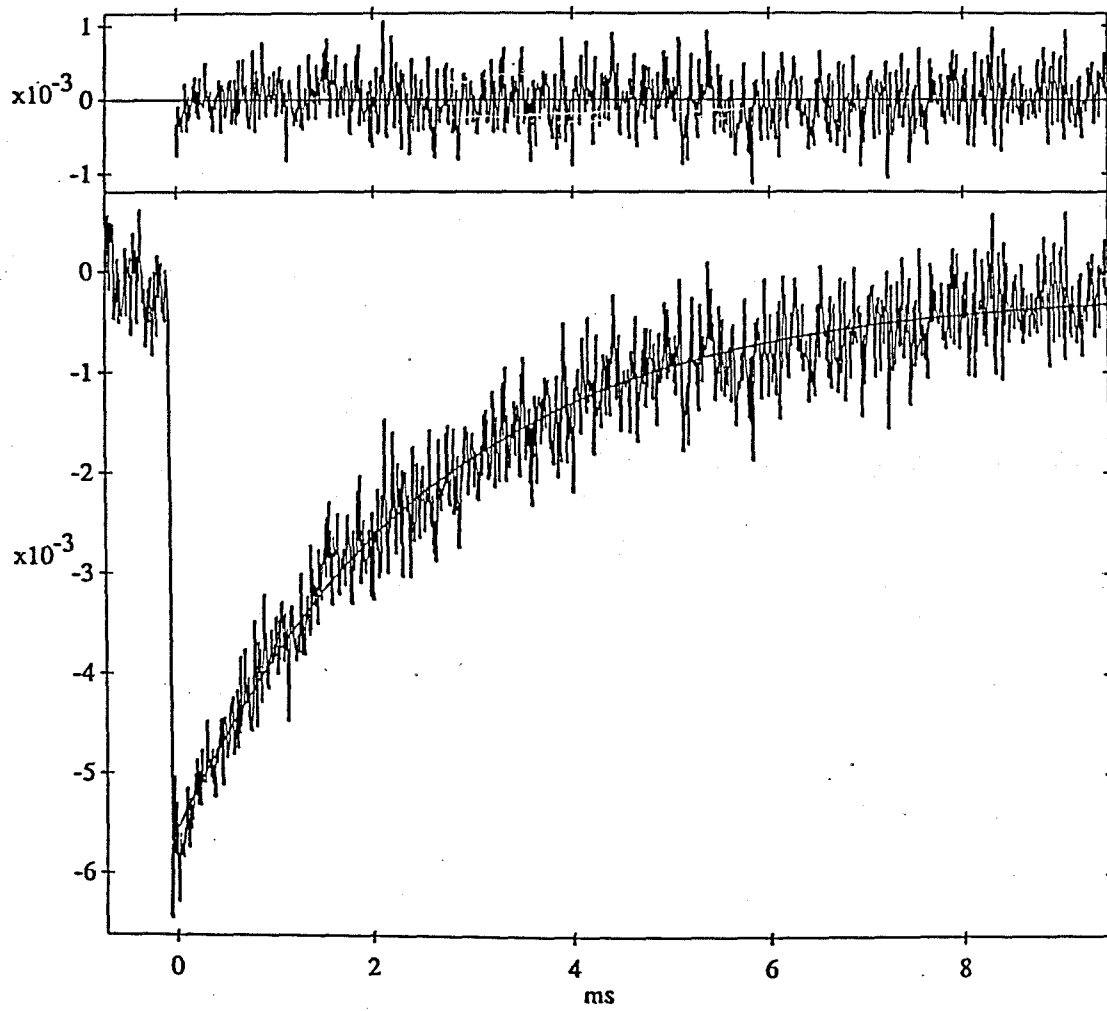


Figure 3.3. Representative kinetics from flash-quench experiments for the H107 mutant. (a) Absorption increase monitored at 625 nm is due to the formation of Cu^{2+} . The concentration was 60 μM . The best fit for this trace gives a rate constant of 920/s. (b) Ru^{2+} formation monitored at 500 nm (protein concentration=60 μM). Best fit line has a value of 990/s. (c) Ru^{2+} formation monitored at 500 nm (30 μM). Average of best fit at this concentration was 600/s. (d) Ru^{2+} formation monitored at 500 nm (protein concentration=20 μM). Average of best fit lines has an average value of 420/s







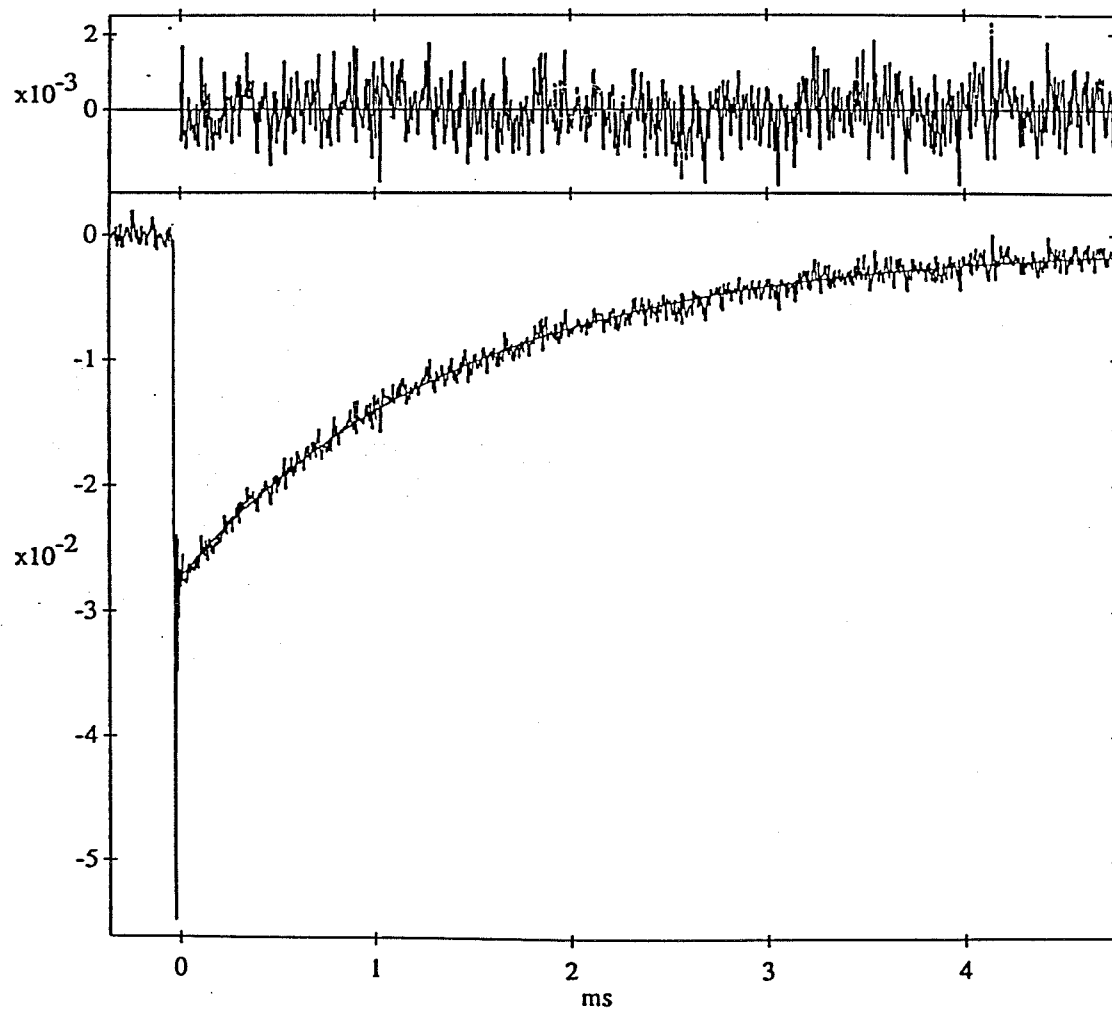


Figure 3.4. Concentration dependence of ET rates for the H107 mutant. The intramolecular ET rate was obtained by extrapolation to infinite dilution. The points represent average values obtained from several measurements at the given concentration.

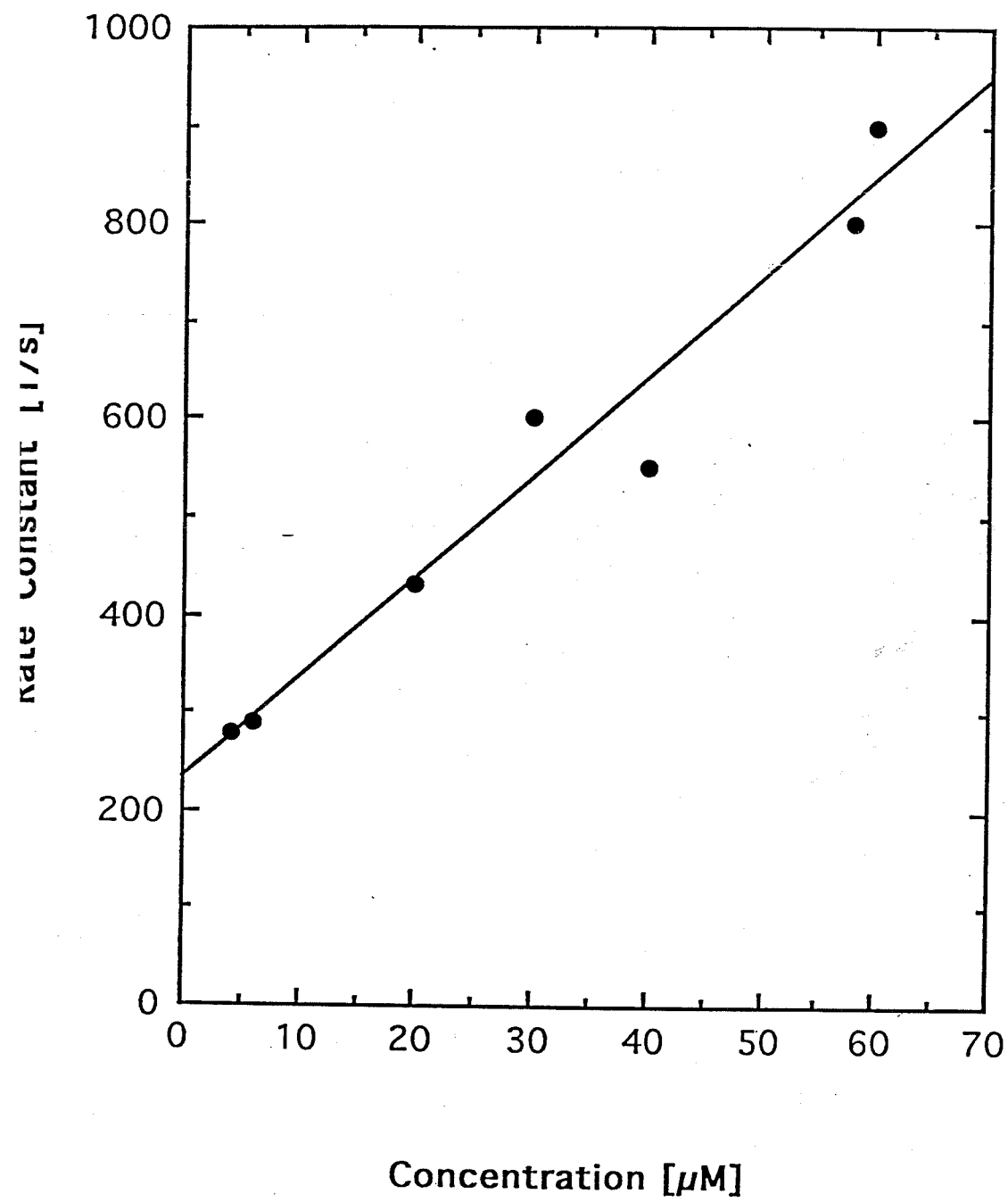


Figure 3.5. The logarithm of the rate constants is plotted versus metal-metal distance in Å. Distances are obtained by modeling. The solid circles represent the rates from the C112 strand; the open circles represent the data from the M121 strand. The fit was constrained to an intercept of 13.

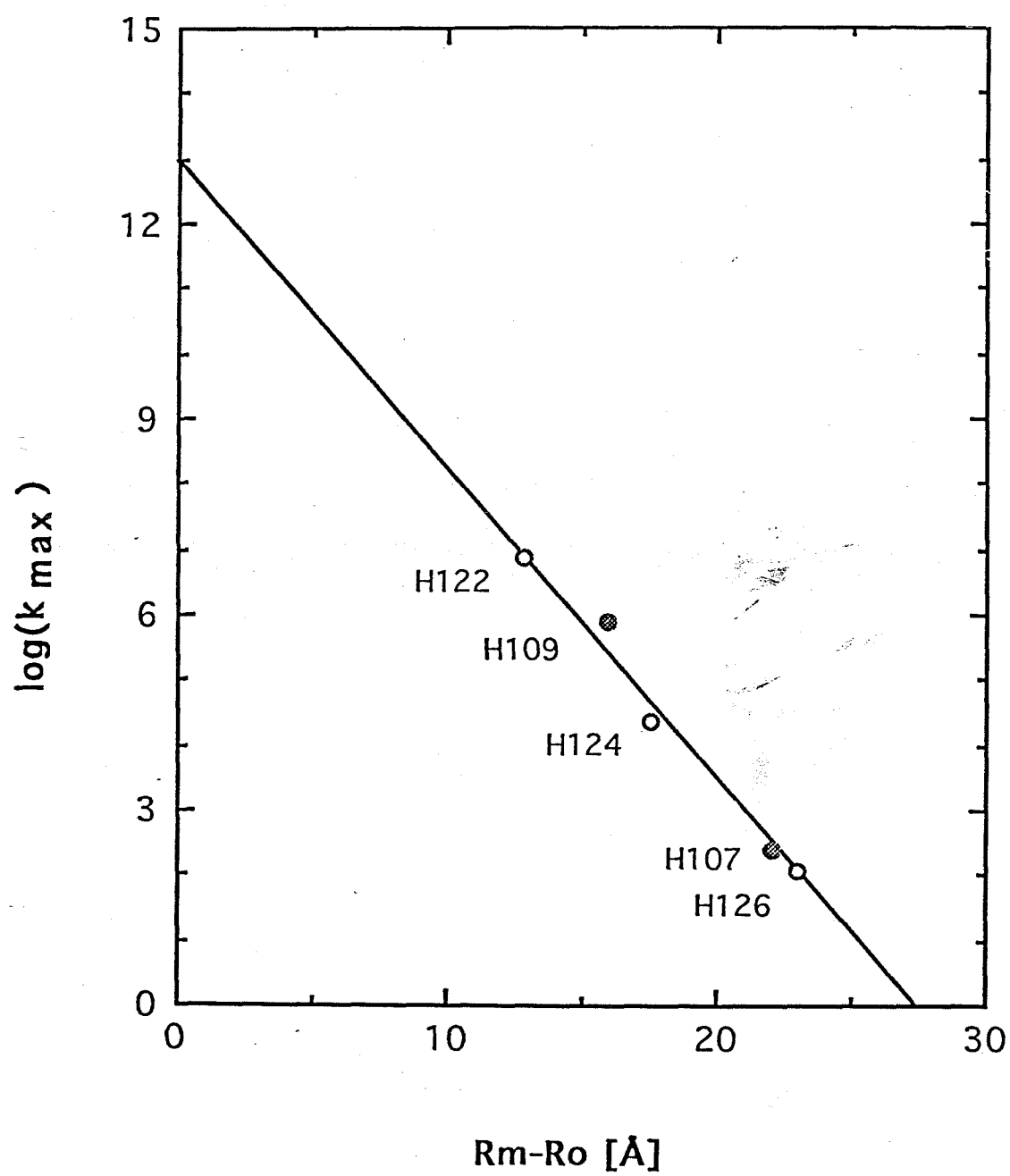


Figure 3.6. The logarithm of the rate constants is plotted versus tunneling length in Å. Distances are obtained by a Beratan-Onuchic pathway search. Empty circles are from tunneling pathway that does not include the ligand-copper interaction. In this representation, the y-intercept would be too low ($2.5 \times 10^{11} \text{ s}^{-1}$).

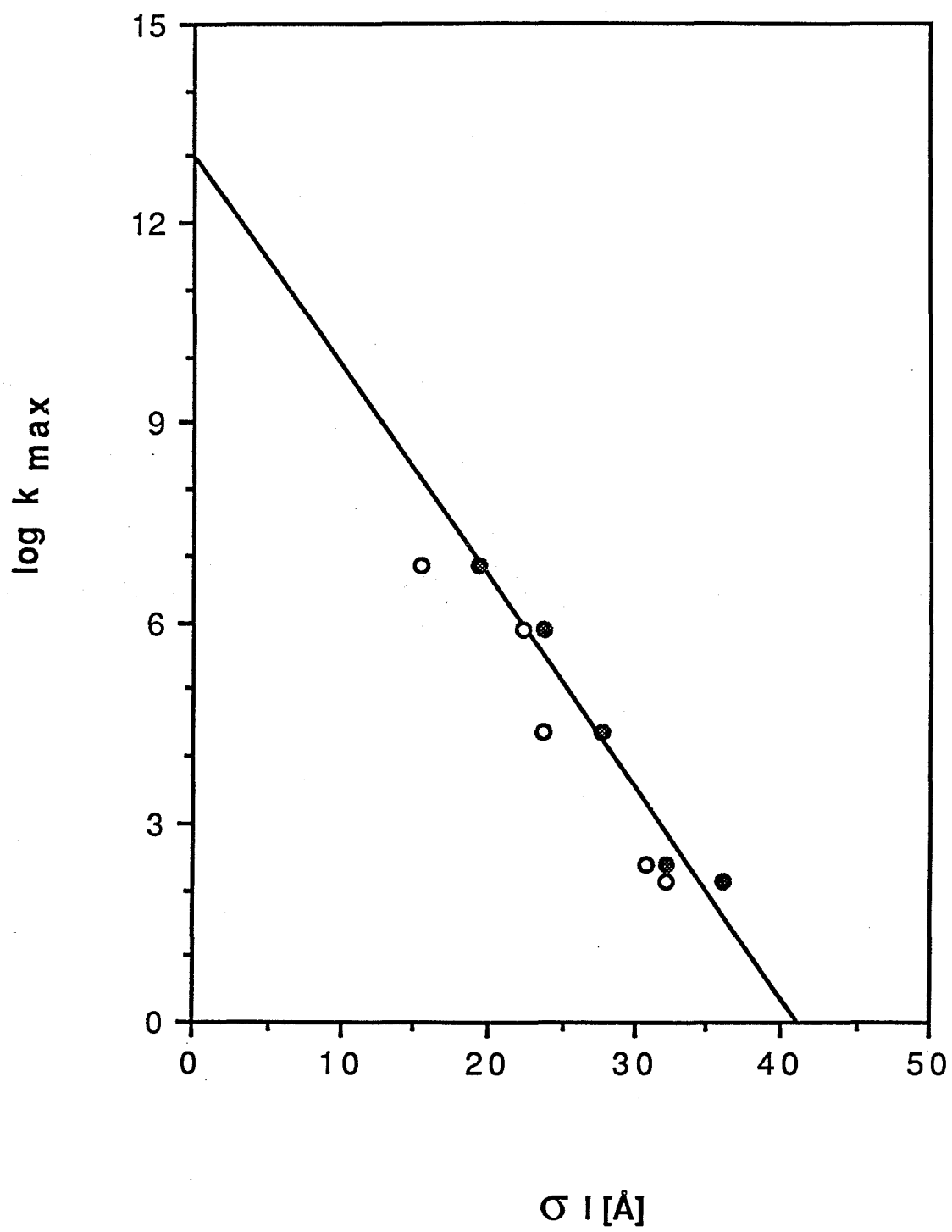


Figure 3.7. Plot of tunneling length as a function of metal-metal separation in ruthenium-modified azurin and cytochrome *c*. Solid line the is the best fit line taken from Fig. 2.10. Solid circles represent cytochrome *c* data, open circles represent azurin data.

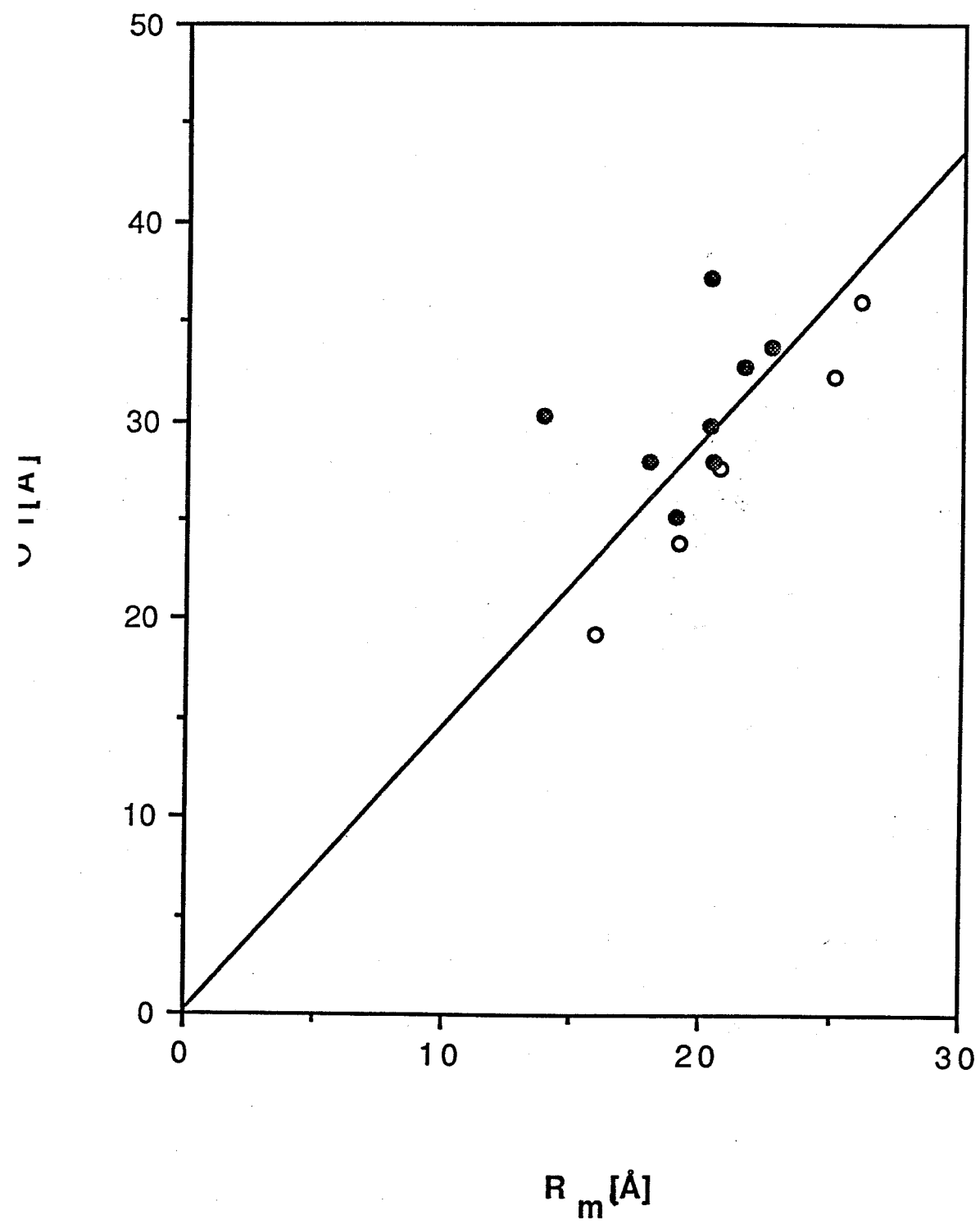
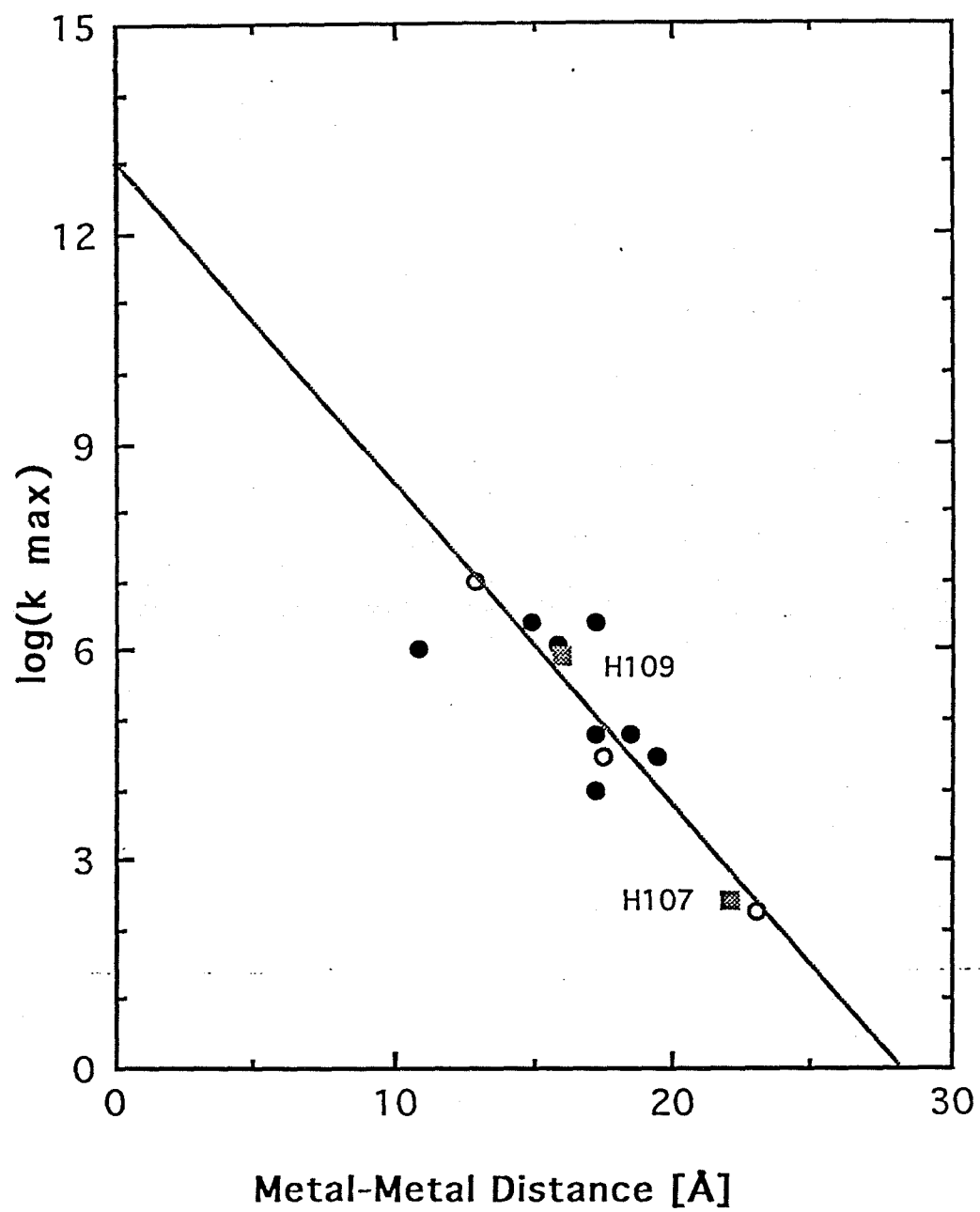


Figure 3.8. Plot of $\log(k_{max})$ vs. metal-metal separation for Ru-modified azurin and cytochrome *c*. Solid line is that predicted from Fig. 2.10. The best fit (excluding H62 and H72) would have a slope of $1.07/\text{\AA}$, which is very close to the predicted $1.06/\text{\AA}$. Solid circles represent cytochrome *c* data; open circles are from the M121-strand in azurin; and squares are from the C112-strand in azurin.



Chapter 4

Effect of the Asn52Ile Mutation on the Redox Potential of Cytochrome c

[This work is described in Langen, R., Brayer, G.D., Berghuis, A.M., McLendon, G., Sherman, F., and Warshel, A. (1992) *J. Mol. Biol.*, **224**, 589-600]

Abstract

Theoretical methods for correlation of sequence changes and redox potential of electron transport proteins are examined using the Asn52Ile mutation in cytochrome *c* as a test case. The first approach is based on using the Protein Dipoles Langevin Dipoles method and the high resolution X-ray structures of the native and the mutant protein. This approach is found to give reliable results where all the solvent molecules are represented by Langevin Dipoles and also when some bound water molecules are represented explicitly. A Free Energy Perturbation method is also found to give reasonable results but at the expense of much more computer time. Finally, an approach, that generates mutant structures from the native structure by molecular dynamics simulation and then uses these configurations in Protein Dipoles Langevin Dipoles calculations is found to give a reasonable prediction of the effect of the mutation on the corresponding redox potential.

INTRODUCTION

Electron transfer plays a central role in biological processes such as respiration and photosynthesis (for reviews see [1-7]). One of the best characterized proteins involved in these processes is cytochrome *c*, a soluble 13 kd protein which serves as a carrier for a single electron in the respiratory chain of plants, animals and procaryotes. Much has been learned about the overall properties that determine its redox potential [8-11], which provides the driving force for electron transfer reactions. However, it is still not clear how all the key residues, which are in close proximity to the heme group contribute to the observed redox value .

In the work presented here we use previously developed microscopic models [12, 13] to explore the effect of a mutation at residue 52 (Asn→Ile) of yeast iso-1 cytochrome *c* on the redox-potential of the protein. This residue is in close proximity to the heme group and is believed to play a key role in the control of redox potential in cytochrome *c*. X-ray crystallographic studies have shown that Asn-52 is hydrogen bonded to the pyrrole ring A propionate group of the heme and to an internally bound water molecule, Wat-166 [14, 15]. The replacement of Asn 52 by Ile causes a loss of this water molecule, so an increase in the redox-potential of the mutant protein might be anticipated. However, the opposite effect has been observed experimentally. This unexpected experimental result can be rationalized by the calculations presented here. Based on the X-ray crystallographic coordinates of the native and the mutant protein in the reduced state we show that the Asn residue destabilizes a positive charge on the heme group more than the Ile residue and that this is the key factor in the change in reduction potential upon mutation. This effect can only be partially compensated by other contributions.

In addition to the specific issue of the effect of Asn 52, we explore in this chapter the feasibility of taking the X-ray coordinates of the native protein as a starting point for generating mutant structures for calculations at the energetics of the mutant protein. Previous experience (e.g. [16]) has indicated that it is possible to reproduce the observed free energy changes associated with mutations while starting from the structure of the native protein. This finding might reflect the fact that the correct trend in energy changes (2nd in particularly in changes of electrostatic energies) can be reproduced by many configurations including local relaxations that do not correspond to the exact structure of the mutant protein. Here we can check this point in a more systematic way since we are fortunate to have high resolution structures for both the native and the mutant protein allowing me to examine the difference in the results obtained with observed and generated mutant respectively. It is found that at least for the Asn 52Ile mutation we can reproduce the observed change in redox potential using both the experimentally derived and the generated mutant structures.

METHODS

The methods for mutation, isolation, and characterization of the native (Asn52) and mutant (Ile52) proteins have been performed and described by Das et al .[17].

Redox potentials for both the native and mutant proteins were determined by analytic potentiometric spectroscopy using an optically transparent electrode as previously described [17]. These potentials were independently confirmed by cyclic voltammetry, using thiobipyridine modified gold electrodes, as described by McLendon and Hill [5].

Both approaches gave identical values for the redox potentials: $E^0=275\text{mV}$ (vs.NHE) for the native (Asn52) protein, and $E^0=220\text{mV}$ (vs.NHE) for the Ile52 mutant.

The X-ray crystallographical determination of yeast iso-1-cytochrome *c* to a resolution of 1.2 Å has previously been published by our collaborator G.D. Brayer [14, 15]. In brief, the methodology used for the N52I mutant can be summarized as follows. A hair-seeding technique [18] produced crystals which were isomorphous with those of the wild-type protein. A complete 1.9 Å data set, consisting of 7910 unique reflections, was collected from one of these crystals on an Enraf-Nonius CAD4-F11 diffractometer. Since the mutant crystals were isomorphous to wild-type, restrained parameter refinement [19] was initiated using the wild-type yeast iso-1-cytochrome *c* structure in which residue 52 was replaced by an alanine as a starting model. During the course of refinement the remaining side chain atoms of residue 52 were unambiguously located and added to the structure. The final model is in good agreement with the experimental diffraction data and has excellent stereo-chemistry, as shown by a rms bond distance deviation from ideality of 0.017 Å and a standard crystallographic R-factor of 18.2 %.

In order to explore the origin of the observed change in redox potential we express the energy of oxidizing the heme group in a given protein (p) as

$$\Delta G(\text{heme}^{\text{red}} \rightarrow \text{heme}^{\text{ox}})^{\text{p}} = \alpha^{(\text{g})} + \Delta G_{\text{sol}}^{\text{p}} \quad (1)$$

where heme^{red} and heme^{ox} designate the heme group in the indicated oxidation states, $\alpha^{(\text{g})}$ is the energy of oxidizing the heme in the gas phase, and $\Delta G_{\text{sol}}^{\text{p}}$ is the change in the free energy of interaction between the heme and its protein plus water surrounding upon

oxidation of the heme, which can be treated as "solvation energy." Using the thermodynamic cycle of Fig. 4.1 we obtain

$$\Delta G_4 - \Delta G_3 = \Delta G_1 - \Delta G_2 = \Delta G_{\text{sol}}^{\text{M}} - \Delta G_{\text{sol}}^{\text{N}} = \Delta G_{\text{elec}}^{\text{M}} - \Delta G_{\text{elec}}^{\text{N}} = \Delta \Delta G_{\text{elec}} \quad (2)$$

where N and M designates the native and the mutant protein, respectively. In this chapter we approximate the solvation energy difference by the corresponding electrostatic energy difference. To realize why this approximation is justified, one has to consider the fact that both ΔG^{N} and ΔG^{M} can be calculated using a two-step thermodynamic pathway that involves first the oxidation of the heme with a fixed protein structure and the protein relaxation process. The first contribution involves only electrostatic energy (except the second-order contribution associated with the minor change in the van der Waals radii of the heme atoms). The second energy contribution, which is associated with the protein relaxation is the well known protein "reorganization energy" which is also an electrostatic energy (see for example [7, 20]). In fact, as long as the protein satisfies the linear response approximation which seems to be the case in many proteins (e.g. [20, 21]), one can evaluate this contribution by simple estimates without using free energy perturbation approaches. In order to estimate the effect of the protein structural relaxation upon oxidation of its cofactor it is useful to consider the thermodynamic cycle of Fig. 4.6 with the conveniently obtained PDL D energies of states (1), (2), (3) and (4) and write:

$$\Delta G_{\text{elec}} = \Delta G(Q^{\text{red}}, \mathbf{r}_{\text{red}}^{\text{p}}, \mathbf{r}_{\text{red}}^{\text{w}}) - \Delta G(Q^{\text{red}}, \mathbf{r}_{\text{ox}}^{\text{p}}, \mathbf{r}_{\text{ox}}^{\text{w}}) = \Delta G_{\text{elec}}^{1 \rightarrow 2} + \Delta G_{\text{elec}}^{2 \rightarrow 3} \quad (3)$$

where \mathbf{r}^{p} , and \mathbf{r}^{w} are the configurations of the protein and the solvent respectively in the indicated oxidation state and the $\Delta G_{\text{elec}}^{1 \rightarrow 2}$, and $\Delta G_{\text{elec}}^{2 \rightarrow 3}$ are defined in Fig. 4.6. In order to estimate $\Delta G_{\text{elec}}^{2 \rightarrow 3}$, note that PDL D calculations with the protein configurations generated

by MD simulations of the oxidized protein can be used to obtain the $\Delta G_{\text{elec}}^{4 \rightarrow 3}$ defined in Fig.4.6. Also note, that $\Delta G_{\text{elec}}^{2 \rightarrow 3}$ can be approximated by

$$\Delta G_{\text{elec}}^{2 \rightarrow 3} = \Delta \Delta G_{\text{Qq}}^{\text{P}} + \Delta \Delta G_{\text{qq}}^{\text{P}} \approx \frac{1}{2} \Delta \Delta G_{\text{Qq}}^{\text{P}} \approx \lambda_{\text{p}}' \quad (4)$$

where $\Delta \Delta G_{\text{Qq}}^{\text{P}}$ is the change in protein-heme interaction due to the relaxation of the protein and where in systems that follow the linear response approximation half of this energy gain is paid for in the increase of the dipole-dipole interaction, $\Delta \Delta G_{\text{qq}}^{\text{P}}$. (Note, that here λ_{p}' is not the standard reorganization energy (the λ of Fig.4.6) since the water molecules are allowed to relax at each state.) Using equations (3), (4) and the fact, that $\Delta G_{\text{elec}}^{4 \rightarrow 3} - \Delta G_{\text{elec}}^{1 \rightarrow 2} \approx \Delta \Delta G_{\text{Qq}}^{\text{P}}$ (where we neglect the change in $\Delta \Delta G_{\text{qq}}^{\text{P}}$ due to the fact that the water molecules are in different configurations in the $4 \rightarrow 3$ and the $1 \rightarrow 2$ transition) we obtain the approximate expression

$$\Delta G_{\text{elec}} \approx \Delta G_{\text{elec}}^{1 \rightarrow 2} + \lambda_{\text{p}}' \approx \frac{1}{2} (\Delta G_{\text{elec}}^{1 \rightarrow 2} + \Delta G_{\text{elec}}^{4 \rightarrow 3}) \quad (5)$$

Since the free energies $\Delta G_{\text{elec}}^{1 \rightarrow 2}$ and $\Delta G_{\text{elec}}^{4 \rightarrow 3}$ can be obtained from PDL D calculations of the reduced and oxidized protein it is possible to estimate the effect of the protein reorganization energy without calculating this quantity. (Note that λ and λ_{p}' can be evaluated by the PDL D approach of Churg *et al.* 1983 or by the FEP of Warshel [20, 22].

In order to evaluate the relevant ΔG 's we used both the Protein Dipoles Langevin dipoles (PDL D) method ([12, 13]) as implemented in the program POLARIS (Warshel & Creighton, 1989) and a free energy perturbation method [22-26] as implemented in the program ENZYMIX [27]. Both methods will be briefly reviewed below.

The PDL D method, which is described in detail elsewhere [12, 13] considers the electrostatic energy of the heme group in the (average) observed X-ray structure as an

approximation for the free energy obtained by considering all the thermally accessible protein and solvent configurations.

The difference in the electrostatic energy $\Delta\Delta G_{\text{elec}}^{\text{P}}$ between the reduced and the oxidized states on the heme is approximated by

$$\Delta\Delta G_{\text{elec}}^{\text{P}} (Q^{\text{red}} \rightarrow Q^{\text{ox}}) \approx \Delta V_{\text{Qq}} + \Delta V_{\text{Q}\alpha} + \Delta G_{\text{Qw}} + \lambda_{\text{p}}' \quad (6)$$

Here ΔV_{Qq} is the energy change due to the electrostatic interactions between the charges on the heme group and the charges on the protein atoms, whereas $\Delta V_{\text{Q}\alpha}$ is the energy change due to the interactions between the charges of the heme group and the induced dipoles of the protein atoms. ΔG_{Qw} is the energy contribution due to the water molecules inside and outside the protein. λ_{p}' is the protein contribution to the reorganization energy associated with the structural relaxation that follows the oxidation process [20].

The charge-charge interaction term ΔV_{Qq} is given (in kcal/mol) by

$$\Delta V_{\text{Qq}} = 332 \sum_{ij} Q_i q_j / r_{ij} \quad (7)$$

where the distances are in Å and the charges are in units of electron charges. The index j refers to the charges on the surrounding protein atoms and the index i refers to the charges on the heme group which were evaluated by the quantum mechanical QCFF/PI method [28]. In order to check the sensitivity of the results to the assumed heme charges

we also used a simplified charge distribution by assigning a charge of 0.25 in the ferric state and 0.0 in the ferrous state to the four heme nitrogen atoms.

The induced dipoles term is given (in kcal/mol) by

$$\Delta V_{Q\alpha} = 166 \sum_j \gamma_j (\xi_j)^2 \quad (8)$$

where γ_j is the polarizability of the j^{th} protein atom (in \AA^3) and ξ_j is the local field at the position of this atom from the residual charges of all other atoms in the system and from all other induced dipoles.

The water molecules inside and around the protein are represented by constructing a cubic grid of 3 \AA spacing (while deleting all points within a van der Waals distance from the protein atoms) and placing a Langevin type dipole on each grid point (see [12, 13] for more details). The corresponding water-heme interaction is given (in kcal/mole) by

$$\Delta G_{Qw} = \Delta G_{QL} + \Delta G_{\text{bulk}} = -166 \sum_k \mu_k \xi_k^0 + \Delta G_{\text{bulk}} \quad (9)$$

Here μ_k is the k^{th} Langevin dipole and ξ_k^0 is the field at the position of this dipole due only to the permanent charges of the system (the Q's and q's) while the polarization of μ_k also reflects the field from the other solvent dipoles [12, 13]. ΔG_{bulk} represents the contribution of the water molecules outside the region represented by Langevin dipoles.

This contribution is evaluated by a continuum approximation [12, 13]. The reorganization energy λ_p' will be neglected in the first stage of the calculations and then considered as a correction term.

In addition to the Langevin dipoles, we included in some calculations explicit water molecules, which are observed in the X-ray structure to be in a close proximity to the heme propionic acids. Such a treatment should allow us to examine whether the effect of relatively fixed water molecules can be reproduced by Langevin dipoles. The explicit water molecules considered are the 121, 166 and 168 water molecules for the wild-type protein and the corresponding 121 and 168 water molecules in the mutant protein. Water molecule 166 was not included in the mutant protein since the mutation resulted in a loss of this molecule.

In parallel to the PDL method we also use a more rigorous and much more computationally intensive FEP method [16, 22-26] which evaluates the energy associated with "charging" the heme. In this approach we use a mapping potential of the form [22]

$$\epsilon_m = (1 - \theta_m) \epsilon_1(Q(H^0)) + \theta_m \epsilon_2(Q(H^+)) \quad (10)$$

where ϵ_1 and ϵ_2 are the force fields of the protein with the reduced and oxidized heme and θ_m is a mapping parameter that controls the heme charges (Q) corresponding to the effective potential ϵ_m (for example when $\theta_m = 0$ we use the charge of the reduced protein).

By a series of molecular dynamics (MD) simulations we change θ_m in n discrete steps and obtain the overall change in free energy by

$$\delta G(\theta_m \rightarrow \theta_{m'}) = - \left(\frac{1}{\beta} \right) \ln [\langle \exp(-(\epsilon_{m'} - \epsilon_m)\beta) \rangle_m] \quad (11)$$

$$\Delta G_{sol}^p = \Delta G(\theta_0 \rightarrow \theta_n) = \sum_{m=0}^{m=n-1} \Delta G(\theta_m \rightarrow \theta_{m+1}) \quad (12)$$

where $\beta = \frac{1}{k_b T}$ and k_b is the Boltzmann constant.

In many cases one is not fortunate enough to have the structure of the mutant protein. Thus, we examine here the possibility to use the X-ray structure of the native protein as the starting point in generating mutant structures for PDLD calculations. The average PDLD energy obtained in this way is used as an estimate of the electrostatic energy of the mutant protein. This type of study is examined here by starting from the native structure but with isoleucine in the position of Asn52 and running trajectories of this mutant system while collecting snapshots for PDLD calculations. Such a procedure will be examined below.

RESULTS

Electrochemical Results

The Asn52Ile replacement results in a 55mV shift in redox potential as shown clearly by the cyclic voltammogram obtained by McLendon's Group (Figure 4.2). It is interesting that although the potential is shifted, there is no apparent effect on the electron transfer kinetics between the protein and electrode. This is indicated by the characteristics of the waveform which indicate a reversible diffusion controlled electron transfer for both systems. A peak to peak separation of 60 mV is indicative of fast reversible electron transfer at the interface; with equivalent areas for the cathodic and anodic peaks, and a peak height which scales as the square root of the scan rate indicates diffusion controlled electrochemistry for this mutant.

Structural effects

The structural modification caused by the Asn52Ile replacement has been determined by X-ray diffraction studies of the wild type and the mutant protein [14, 15, 29, 30]. The key findings are summarized in figures 4.3 and 4.4. In brief, this mutation causes no appreciable shifts in the main chain backbone of the proteins. Instead, structural dislocation is localized at the point of mutation. As shown in figure 4.4b the bulky Ile52 not only replaces Asn52, but also displaces a key internal water molecule (Wat-166) which normally forms a hydrogen bond network with Asn52, Tyr67, and Thr-78. A secondary compensating shift occurs at Tyr67 to fill the residual void left by the displaced water, with formation of an additional hydrogen bond to Thr78.

PDL D calculations

The PDL D method was used to evaluate the effect of the Asn 52Ile mutation on the redox potential of the protein. The calculations were done with the relevant X-ray structures of the native and mutant protein. Tables 1 and 2 summarize the corresponding results, which

were obtained with different assumptions about the charges of the system and the inclusion of some explicit water molecules. The sensitivity of computer simulations to the parameters and the conditions used provides probably the best estimate of the error range associated with the corresponding conclusions. As can be seen from this table, the contributions to ΔG_{elec} from ΔV_{Qq} , $\Delta V_{\text{Q}\alpha}$, and ΔG_{Qw} varies significantly, depending on the choice of charges and the use of explicit water molecules. For example, assigning a negative charge to each propionic acid decreases ΔV_{Qq} by more than 80 kcal/mole, whereas the bulk energy increases by approximately 49 kcal and $\Delta V_{\text{Q}\alpha}$ increases by 20 to 30 kcal for both native and mutant protein. Although these are very substantial changes they follow the large compensating dielectric effect expected from charge-charge interactions in solvated proteins [12, 13, 31] and the difference between the native and mutant cytochromes remains almost constant (relative to the changes in the various contributions). The average value of the calculated $\Delta\Delta G$ in the presence of the ionized propionic acid is -2.26 kcal/mol. This is equivalent to a 98 mV difference in redox potential as compared to the experimentally determined value of 55 ± 5 mV. This overestimation is due to the fact that the calculations are based on the X-ray structure of the reduced protein, and do not reflect the structural relaxation that follows the oxidation process. That is, as was mentioned in the previous section, one should consider in calculating ΔG_{elec} the reorganization energy (λ) associated with the structural changes between the reduced and the oxidized cytochrome *c* [20]. Although this effect can be evaluated quite rigorously we choose here to use the approximation ,

$$\Delta G_{\text{elec}} \approx \Delta G_{\text{elec}}^{1 \rightarrow 2} + \lambda_p' \approx \frac{1}{2} (\Delta G_{\text{elec}}^{1 \rightarrow 2} + \Delta G_{\text{elec}}^{4 \rightarrow 3}) \quad (13)$$

where $\Delta G_{\text{elec}}^{1 \rightarrow 2}$ and $\Delta G_{\text{elec}}^{4 \rightarrow 3}$ represent the PDL D electrostatic energies obtained respectively with the structures of the reduced and oxidized protein (see Fig.4.6 and eq.3-

6). Here λ_p' is the protein contribution to the reorganization energy. This contribution is expected to be larger in the native protein since Asn is more polar than Ile and expected to undergo larger relaxation upon oxidation. In fact using MD to generate structures of the oxidized protein we find, that $(\Delta G_{elec}^{4 \rightarrow 3})_N \approx (\Delta G_{elec}^{4 \rightarrow 3})_M$. Thus, $\Delta \Delta G_{elec} \approx -1.2$ kcal/mol or 51.96 mV in an excellent agreement with the observed value.

A closer look at the individual contributions leading to ΔG_{elec} shows that ΔV_{Qq} , which basically represents the interaction between the heme and the residual charges (permanent dipoles) of the protein is consistently more negative by 4 to 7 kcal/mol in the mutant than in the native protein. This means that the residual charges (permanent dipoles) of the mutant protein are better at stabilizing the positive charge on the heme group. Although $\Delta V_{Q\alpha}$ and ΔG_{Qw} compensate partially for the effect of ΔV_{Qq} the overall effect retains the trend of ΔV_{Qq} . Thus, it appears that the protein residues (rather than the solvent) are responsible for the fact that the mutant protein stabilizes the oxidized heme better than the native protein.

A comparison of the contributions from every single residue of the native and the mutant protein to the heme electrostatic potential has indicated, that all the unaltered residues (all residues except residue 52) have very similar effects in the native and the mutant protein.

The electrostatic contributions to $\Delta \Delta G$ from residue 52 were examined by several approaches which are summarized in Table 3. Part A of this table gives the contributions obtained using the PDL method based on the X-ray structure of the native and mutant protein and two rather different charge sets. Both calculations gave very similar results where the Asn residue destabilizes the oxidized heme by nearly 3 kcal/mol more than the Ile residue. Part B of the table summarizes the results obtained by a series of PDL calculations using the structures generated by MD simulation with the conditions indicated in the table. The results obtained by this molecular dynamics approach are very

similar to those obtained using the observed X-ray structure. The reason for the electrostatic effect of Asn52 became clear upon inspection of Fig. 4.4. As seen from this figure and from simple electrostatic considerations the Asn residue stabilizes the negatively charged propionates by its Amide dipole. This same dipole, however, destabilizes a positive charge on the heme and makes it harder to oxidize the heme.

The present calculations allow one to examine the contribution of the solvent molecules to $\Delta\Delta G$. This contribution is of significant interest since the Asn52Ile mutant involves a displacement of a water molecule by the bulky isoleucine group, where one could expect this effect to be associated with destabilization of the oxidized heme. However, this water molecule (in both the explicit and the Langevin representations) does not seem to be the dominant factor for the change in redox potential in the mutant protein. This conclusion is of course, not a general one. For example in the case of the Phe82Ser mutation water molecules seem to play a crucial role in changing the redox potential [14, 15, 29].

Nevertheless, while the effect of the solvent can be different in different cases, it is encouraging to note that the electrostatic contribution associated with the explicit water molecules can be represented by Langevin dipole even when we deal with rather fixed water molecules. This, however, might not be the case when one deals with an isolated solvent molecule that interacts with charges that are more localized than the heme charges.

Another issue that has been examined by the present approach is the effect of the charges of the propanoic acids on the redox potential of the heme. The calculations obtained without explicit water molecules (Table 1) produce only small effects (at most 3 kcal/mol) for the change in redox potential upon neutralization of the charges of the propionic acids. This represents an almost complete compensation for the enormous gas

phase interaction (the ΔV_{Qq}). The calculations obtained with some explicit water molecules (Table 2) give a less complete compensation. This reflects the fact that the explicit water molecules were not allowed to reorient upon neutralization of the propionic acids while the Langevin dipoles were free to reorient for any new charge distribution. The predicted effect of the propionic charge on the overall ΔG_{elec} is much smaller than the effect expected from charge-charge interactions in a medium of low dielectric constant demonstrating, that the protein should not be treated as a non polar medium. This prediction is consistent with recent experimental studies by Funk *et al.* [32], who observed that charging the propionic acids in cytochrome b does not contribute significantly to the redox potential of this protein. It is also in agreement with previous calculations [11] that predicted a large effective dielectric constant for the interaction between the heme and the propionic acids despite their location in the protein interior.

Free energy perturbation calculations

In order to further examine the validity of the PDL D calculations and the validity of the conclusions we also performed FEP calculations for the native and the mutant protein. This method represents the water molecules by an explicit all atom representation rather than by Langevin dipoles and takes the protein reorganization energy into account [24]. The calculations, which are much more complex than the PDL D calculations involved negatively charged propionic acids and QCFF/PI heme charges. The heme van der Waals radii were assumed to be the same in the reduced and oxidized states of the heme. The initial position of the water molecules inside the protein cavities were generated by a search procedure except water molecules 121 and 168 whose initial positions were taken from the X-ray structure. The calculations which are summarized in table 4 gave a $\Delta\Delta G$ of -2.82 kcal/mol which somewhat overestimates the observed $\Delta\Delta G$. Thus, despite whatever convergence and/or boundary condition problems the free energy perturbation method may have [31], we were able to reproduce the observed trend, with the mutated protein having a lower redox potential than the native protein. As an additional check of

the FEP calculations we examined the dependence of the calculated ΔG_{sol} and $\Delta\Delta G$ on the initial coordinates used. (Note, that also in the present case ΔG_{sol} can be identified as an electrostatic energy since the effect of the change in the heme nonbonded parameters upon oxidation is expected to be negligible.) The corresponding results are summarized in Table 4. As seen from the table, the $\Delta\Delta G$ value obtained from the mutant structure generated by MD from the native structure (see below) is smaller than that obtained by starting from the X-ray structure of the mutant protein. Although the $\Delta\Delta G$ obtained from the generated mutant is in better agreement with the experimental $\Delta\Delta G$. This difference probably reflects an incomplete convergence for the MD generated structures and can serve to establish an error bar on the results of the FEP calculations

Generating artificial mutant protein structures for PDL D calculations

As stated in the methods section, one of the purposes of this work is to compare the energetics obtained with the X-ray structure of the mutant to that obtained by generating mutant structures. This comparative study should tell us how reliable this approach would be. Obviously, such a comparative test can only be done when the structure of the mutant is known and it would become more and more relevant with an increasing number of test cases. We started from the structure of the native protein, and then exchanged the Asn 52 coordinates for isoleucine and relaxed this structure using the ENZYMIX program. Coordinate sets were obtained in equal time steps during the relaxation process and used in PDL D calculations. The calculated energies for the proteins were then averaged as depicted in Fig.4.5. Applying the same modeling conditions to the native, mutant, and generated mutant proteins, yielded quite stable average ΔG 's for all three structures after only a few picoseconds.

As shown in Fig. 4.5, the $\Delta\Delta G$ from the mutant and the native reference structure reaches a final value of -2.95 kcal/mol, which is reduced to approximately -1.5 kcal/mol when the reorganization energy is taken into account. This is in a reasonable agreement with the

experimentally observed value of -1.27 kcal/mol. The $\Delta\Delta G$ between the native and the generated mutant protein reaches a final value of -1.39 kcal/mole, which is reduced to -0.7 kcal/mol after consideration of the reorganization energy. This is smaller than the above value, probably reflecting an incomplete convergence.

CONCLUSIONS

The present study is not aimed at the challenging issue of reproducing the mutant structures. What we try here is to examine what does it take (in terms of computer time) to reproduce the energetics of a mutant protein starting from the structure of the native protein. Our studies and that of others [16, 25, 26] have demonstrated that one can reproduce the energetics of mutant enzymes starting from the native structure. Similarly, the present study reproduced the observed effect of the mutation on the redox potential. We suspect that these successes reflect the fact that the physics of electrostatic response is rather simple so that there are many configurations with similar energetics and some of them would be found rather easily by introducing the system to a given perturbation. In this respect, it is important to recognize the fact that the energetics of an ion in water can be reproduced by billions of different solvent configurations and that computer simulations are unable to span all these configurations but can easily produce the correct solvation energy. Thus, it is probably much simpler to reproduce the electrostatic response of the protein to the deletion of the Asn dipole than to reproduce the exact mutant structure. What is done here is the demonstration that one can estimate the effect of mutations on redox in much less time than needed for reproducing the overall observed structural change. This finding is supported by a similar finding for the Phe82 Ser mutant (Langen et al. to be published), and of course by the above quoted studies of mutations in enzymes. In this respect, what is new in the present study is the availability of high

resolution mutant structure that allows one to assess the error range of the presented approach by comparing the results obtained with the generated structures to that obtained with the actual average structure. That is, what is done here in comparison of energetics in different structures rather than comparison of structures. The generality of these findings will require, of course, studies of more systems but it is clearly unrelated to the fact that some readers might not believe that changes in electrostatic energies are simpler to determine than changes in structures.

The Asn52Ile mutation is of interest for several reasons. First, the effect of the Asn52Ile mutation is difficult to rationalize based on structural information alone. That is, since the Ile residue is hydrophobic, in addition to displacing a bound water molecule, it would be reasonable to assume that this mutation will increase the redox potential [29], while the opposite effect is observed experimentally. Secondly, this is one of the rather few cases where we have high quality X-ray structures of the native and mutant protein [14, 15, 30]. Finally, the structure of bound water molecules is well resolved in this system so one may compare calculations with explicit bound water to those where the bound water are replaced by Langevin dipoles.

The present study has reproduced the unexpected effect of the Asn 52Ile mutation by all the methods used. This indicates that the change in redox potential is associated in the present case with the loss of the repulsive electrostatic interaction between the Asn 52 dipole and the positively charged heme. Apparently this interaction is not completely compensated for by the rearrangement of the rest of the protein and water system. Obtaining correctly the "dielectric effect" associated with the compensation of changes in charge-charge or charge-dipole interactions is in fact the primary challenge of any calculation of redox control by proteins [11, 31].

The accuracy of the present work is reasonable giving values with a range from 15.8 to 64 mV for our various PDL D calculations and 47.72 to 122 mV for our FEP calculations compared to an experimental difference of 55mV. It should be pointed out in this respect that we have used here microscopic calculations, trying to obtain the compensation of large numbers and thus the precision of our calculations is equal to their accuracy (see discussion in ref. [27]). It is significant that the PDL D calculations that take the X-ray structures at their face value give similar results with and without explicit water (provided that the explicit water are replaced by Langevin dipoles). This indicates that one can use high quality X-ray structures to estimate changes in redox potentials upon mutations even if the water molecules are not resolved.

As stated above, the main focus of this work was not to improve current structure determination techniques, but to find a shortcut that provides an estimate of the mutant energetics starting from a known native structure. Such a procedure was already used in FEP studies of enzymes and the present case allows us to compare the energetics obtained with the generated structures to that obtained using the observed mutant structure and to assess the error in terms of energy (rather than structure) associated with the use of the generated mutant and with the intentionally short simulation time.

Some readers might be concerned about a possible contradiction between the procedure used in averaging the PDL D energies over the generated mutant structures and the study of Wendoloski and Matthew [33] who reported major fluctuations in the electrostatic energy during MD simulations. It is true that the fluctuation of the electrostatic energy should be handled consistently as was done repeatedly in our studies of the dynamics of chemical processes in proteins [7, 12], and in all FEP calculations. It is assumed there that the electrostatic energy averaged over the protein fluctuation should be approximated by the electrostatic energy of the average protein structures. This crucial issue which is

perhaps the most relevant issue with regards to the validity of using the static protein structure, has not been examined by Wendolski and Matthew, but is supported by our extensive studies.

In considering the protein contributions to redox potential, one might wonder about the structural reorganization that occurs upon oxidation or about dynamical effects such as the increase in protein flexibility upon oxidation. It is important to realize, however, that the structural reorganization of the protein is represented by the reorganization energy (the 1 of eq. (6)), which is an electrostatic contribution. This contribution does depend on non electrostatic non bonded interactions but, as long as the protein follows the linear response approximation, one finds that about half of the change in the charge-dipole energy (V_{QP}) is invested on the dipole-dipole and non bonded energies and the new contributions of about half V_{QP} is what one is expected from macroscopic electrostatic considerations (see ref. [20]). As to the increase in the cytochrome flexibility upon oxidation; this effect seems to reflect the fact that the electrostatic energy of the oxidized heme is lower by about 7 kcal. mol in water than in the site of the heme in the protein [11]. This fact changes the dynamics of the protein but is not a result of the dynamics. More specifically free energies reflect an average over phase space but not the dynamics of the given fluctuations.

The present study confirms the conclusions from previous studies [11, 32] with regards to the interaction between the charged propionic acid and the heme. It is found that despite the fact that the propionate groups are partially shielded from the solvent their effect on the energetics of the heme is surprisingly small (much smaller than what would be obtained using coulomb's law and a small dielectric constant). Thus it is encouraging to find out that microscopic calculations without an assumed dielectric constant can produce the proper large compensation to the heme-propionate interaction (see ref. [11] for discussion).

Some concern may be expressed about the fact that this work has not considered the effect of the ionic strength and the ionized surface groups. However, as has been demonstrated repeatedly (e.g. Churg et. al 1983), the effect of the surface charges on the redox potential area second-order effects especially when one considers small structural changes such as those induced by the mutations (see Appendix 1 of ref. [20] for discussion). With a negligible effect of surface groups, one would expect an even smaller effect of the ionic strength. This point is indeed proven experimentally. Thus despite the interest in the effect of ionic strength on redox potentials, the importance of this effect is much less than that of the localized protein dipole such as Asn 52.

Although the effect of the mutation was known a priori in the present case this study is much more than an exercise in computer modeling. That is, in many cases when the effect of mutations is unknown it is crucial to establish the error range of the given simulation method. Perhaps the best example is the study of bacterial reaction centers where the mechanisms of the primary charge separation is not clear (it is not clear whether this event involves a sequential mechanism or a supercharge mechanism [7]. In this system one finds large effects of mutations on the observed rate constant [7, 34] and these effects can be used to elucidate the nature of the primary event (see for example ref. [35]). Unfortunately each mutation involves a change of the redox potential of the relevant donor or acceptor which cannot be determined experimentally. A computer modeling approach that can estimate the effect of the mutation on the relevant redox to within say 1 kcal/mol will probably make genetic engineering the most effective tool in checking the nature of bacterial photosynthesis. On the other hand, if the best theoretical estimates are found to have a 3 kcal/mole error range it means that the results of mutation experiments cannot be used in a unique interpretation of the primary event.

Finally it is important to note that the presented computer aided technique might be useful in being less time consuming than the conventional approach of obtaining a mutant

protein by site directed mutagenesis, cloning and purification of the mutant and finally X-ray crystallographic analysis. In principle, this technique might enable one to screen the effect of various mutations on the thermodynamic properties (like ΔG) of proteins in a relatively short time frame, thus facilitating understanding of the relationship between structure and function in redox proteins.

REFERENCES

1. Adman, E.T., *Biochem. Biophys. Acta* **549**, 107 (1979)
2. Salemme, F.R., *Ann. Rev. Biochemistry* **46**, 299 (1977)
3. Mayo, S.L., Ellis, W.R., Crutchley, R.J., and Gray, H.B., *Science* **233**, 948 (1986)
4. Gray, H.B., and Malmström, B.G., *Biochemistry* **28**, 7499 (1989)
5. McLendon, *Acc. Chem. Res.* **21**, 160 (1988)
6. Louie, G.V., and Brayer, G.D., *J. Mol. Biol.* **210**, 313 (1989)
7. Warshel, A., and Parson, W.W., *Ann. Rev. Phys. Chem.* **42**, 279 (1991)
8. Marchon, J.C., Mahiko, T., and Reed, C.: *Electron Transport and Oxygen Utilization*, New York: Elsevier North Holland; 1982.
9. Rafael, A., and Gray, H.B., *J. Am. Chem. Soc.* **113**, 1038 (1991)
10. Rafael, A., and Gray, H.B., *Proteins* **6**, 338 (1989)
11. Churg, A.K., and Warshel, A., *Biochemistry* **25**, 1675 (1986)
12. Warshel, A., and Russel, S., *Quart. Rev. Biophys.* **17**, 283 (1984)
13. Russel, S., and Warshel, A., *J. Mol. Biol.* **185**, 389 (1985)
14. Louie, G.V., Hutcheon, W.L.B., and Brayer, G.D., *J. Mol. Biol.* **199**, 295 (1988)
15. Louie, G.V., and Brayer, G.D., *J. Mol. Biol.* **214**, 527 (1990)
16. Warshel, A., Sussman, F., and Hwang, J.K., *J. Mol. Biol.* **201**, 139 (1988)

17. Das, G., Hickey, D., McLendon, D., McLendon, G., and Sherman, F., *Proc. Natl. Acad. Sci.* **86**, 496 (1989)
18. Leung, C.J., Nall, B.T., and Brayer, G.D., *J. Mol. Biol.* **206**, 783 (1989)
19. Hendrickson, W.A., and Konnert, J.: *Biomolecular Structure, Function, Conformation and Evolution*, Oxford: Pergammon Press; 1981.
20. Churg, A.K., Weiss, R.M., Warshel, A., and Takano, T., *J. Phys. Chem.* **87**, 1683 (1983)
21. Yodav, A., Jackson, R.M., Holbrook, J.J., and Warshel, A., *J. Am. Chem. Soc.* **113**, 4800 (1991)
22. Warshel, A., *J. Phys. Chem.* **86**, 2218 (1982)
23. Postma, J.P.M., Berendson, H.J.C., and Haak, J.R., *Faraday Symp. Chem. Soc.* **17**, 55 (1982)
24. Warshel, A., Sussman, F., and King, G., *Biochemistry* **25**, 8368 (1986)
25. Sing, U.C., Brown, F.K., Bash, P.A., and Kollman, P.A., *J. Am. Chem. Soc.* **108**, 1607 (1987)
26. Wong, C.F., and McCammon, J.A., *J. Am. Chem. Soc.* **108**, 3830 (1986)
27. Warshel, A., Chu, Z.T., and Parson, W.W., *Science* **246**, 112 (1989)
28. Warshel, A., and Lappicierella, L., *J. Am. Chem. Soc.* **103**, 4664 (1981)
29. Louie, G.V., Pielak, G.J., Smith, M., and Brayer, G.D., *Biochemistry* **27**, 7870 (1988)
30. Hickey, D.R., Berghuis, A.M., Lafond, G., Jaeger, J., Cardillo, T., McLendon, D., Das, G., McLendon, G., Sherman, F., and Brayer, G.D., *J. Biol. Chem.* **266**, 1686 (1991)

31. Warshel, A., and Aqvist, J., *Ann. Rev. Biophys. Biophys. Chem.* **20**, 267 (1991)
32. Funk, W.D., Lo, T.P., Mauk, M.R., Brayer, G.D., MacGillivray, R.T.A., and Mauk, A.G., *Biochemistry* **29**, 5500 (1990)
33. Wendolowski, J.J., and Mathews, J.B., *Proteins: Structure, Function and Genetics* **5**, 313 (1989)
34. Coleman, W.J., and Youvan, D.C., *Ann. Rev. Biophys. Chem.* **18**, 333 (1990)
35. Parson, W.W., Chu, Z.T., and Warshel, A., *Biochem. Biophys. Acta* **1017**, 251 (1990)

Figure 4.1. A thermodynamic cycle that relates the redox potentials of the native and the mutant cytochrome *c*. Q^{red} and Q^{ox} designate the charges of the reduced and oxidized heme, respectively.

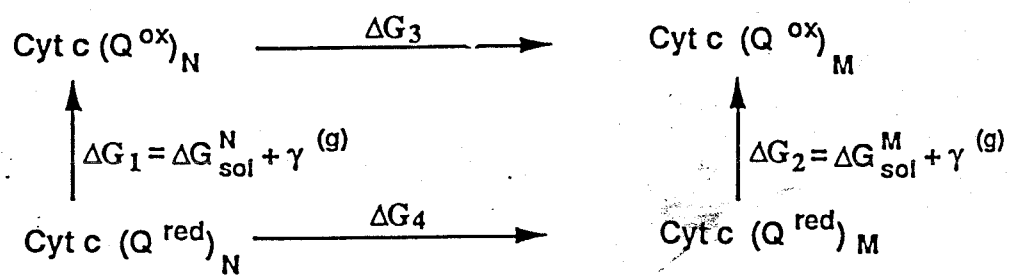


Figure 4.2. Cyclic voltammogram of Asn52Ile cytochrome *c* at a 4,4' bipyridyl disulfide modified Au electrode; scan rate 5 mV/s, pH 7 (Hepes); 100 μ M KCl.

Current

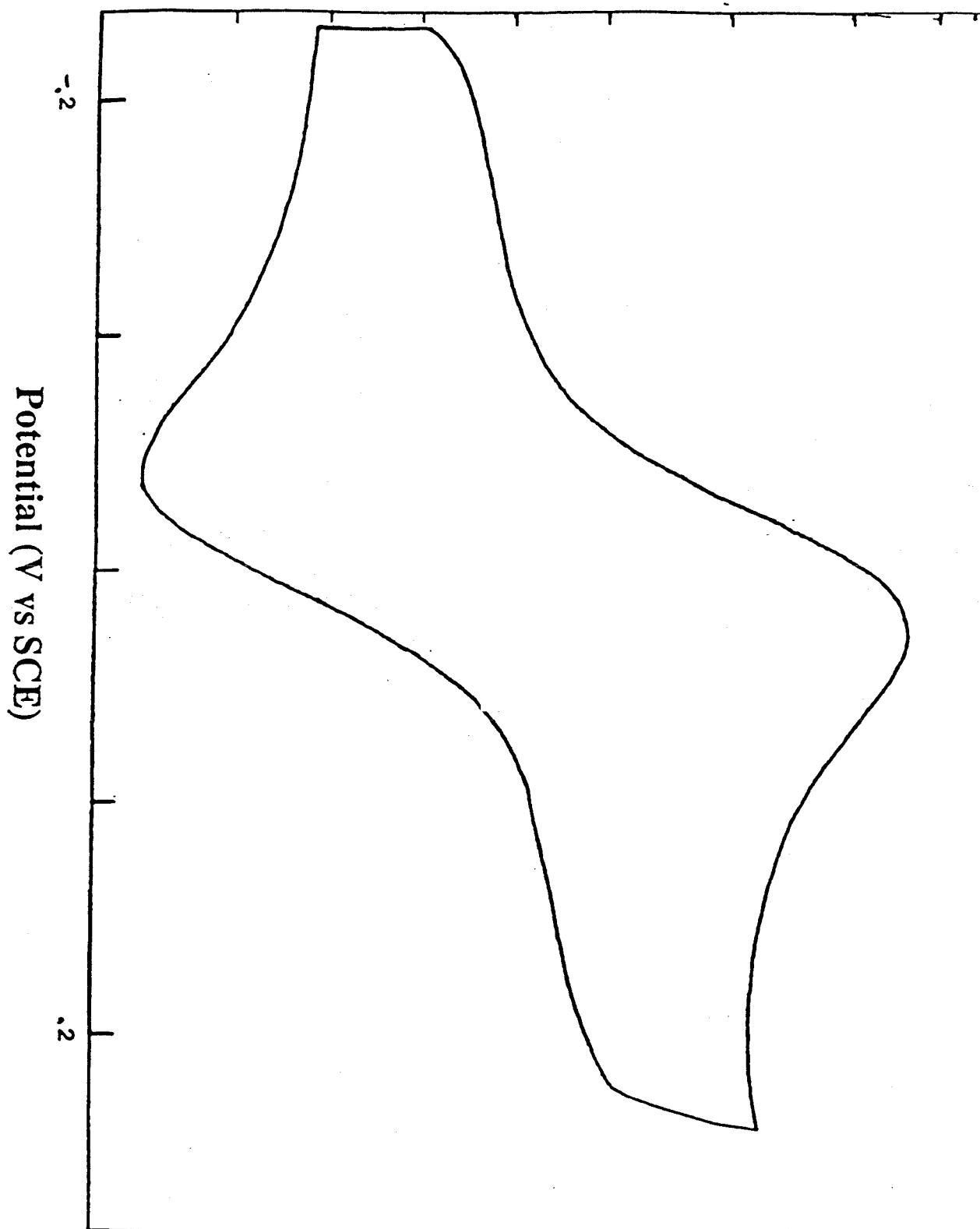


Figure 4.3. A ribbon drawing of the polypeptide backbone chain of yeast iso-1-cytochrome *c*. Also shown using a ball and stick format, with darker shading, is the positioning of the central heme group, which is viewed edge-on in this orientation. The position of the two heme ligands Met-80 and His-18 are also drawn (light shading), along with the side chain of Asn-52 (dark shading). The two heme propionates are located to either side of the side chain of Asn-52.

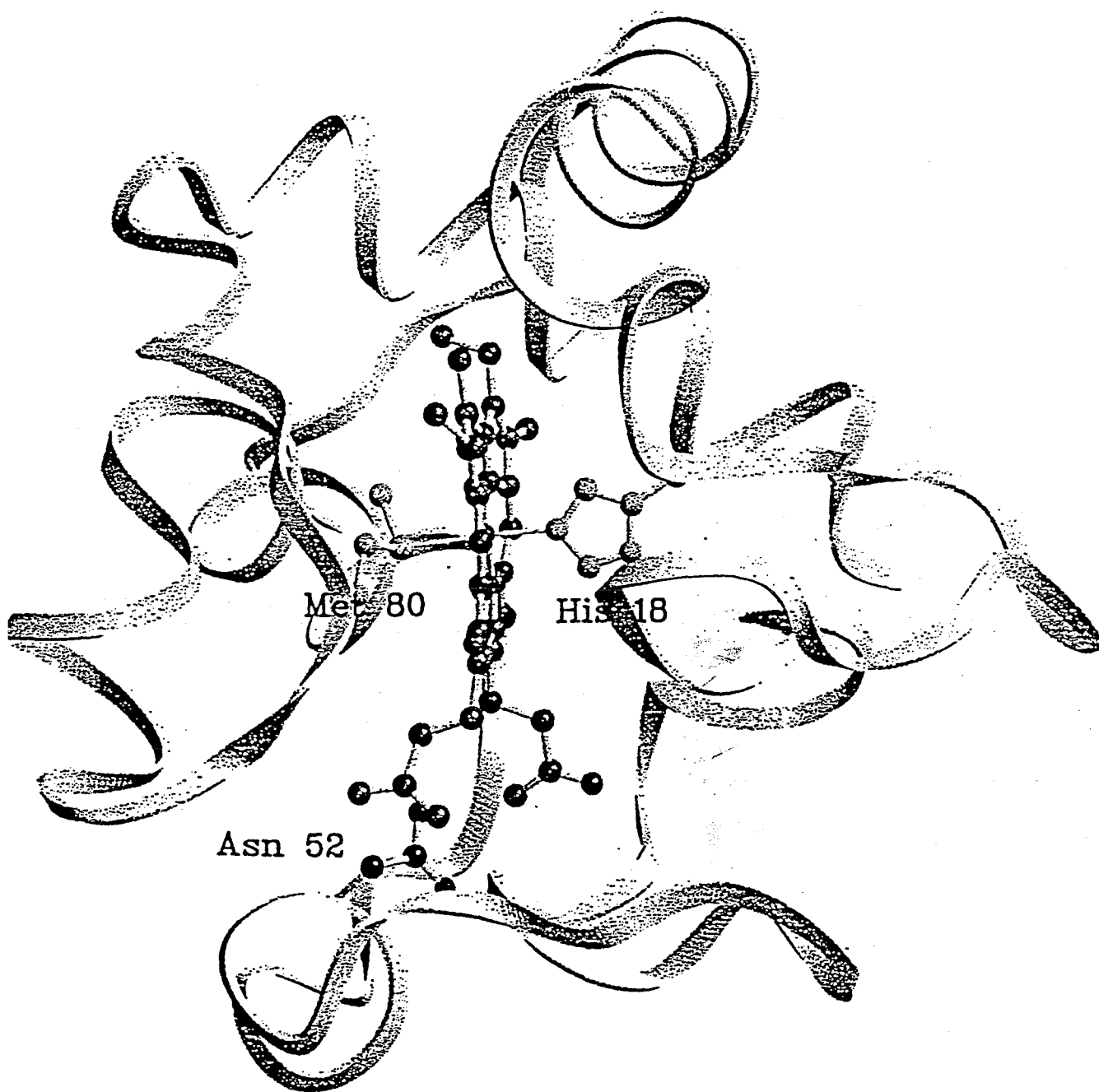


Figure 4.4. Drawings of the immediate vicinity of a) Asn-52 in wild-type yeast iso-1 cytochrome *c* and of b) Ile-52 in the Asn52Ile mutant protein. Only the side chains of nearby residues are drawn, along with observed hydrogen bonds (dashed lines). Heme group atoms are drawn with darker shading and the pyrrole ring A propionate group has been labelled. An internal water molecule, Wat-166, is shown as a larger sphere in the wild-type protein, but is not present in the Asn52Ile mutant. The side chain of Asn-52, through the orientation of its dipole destabilizes the oxidized form of the heme group. Removal of this interaction in the mutant Asn52Ile protein results in a smaller destabilization of the the oxidized state and a corresponding decrease in observed reduction potential.

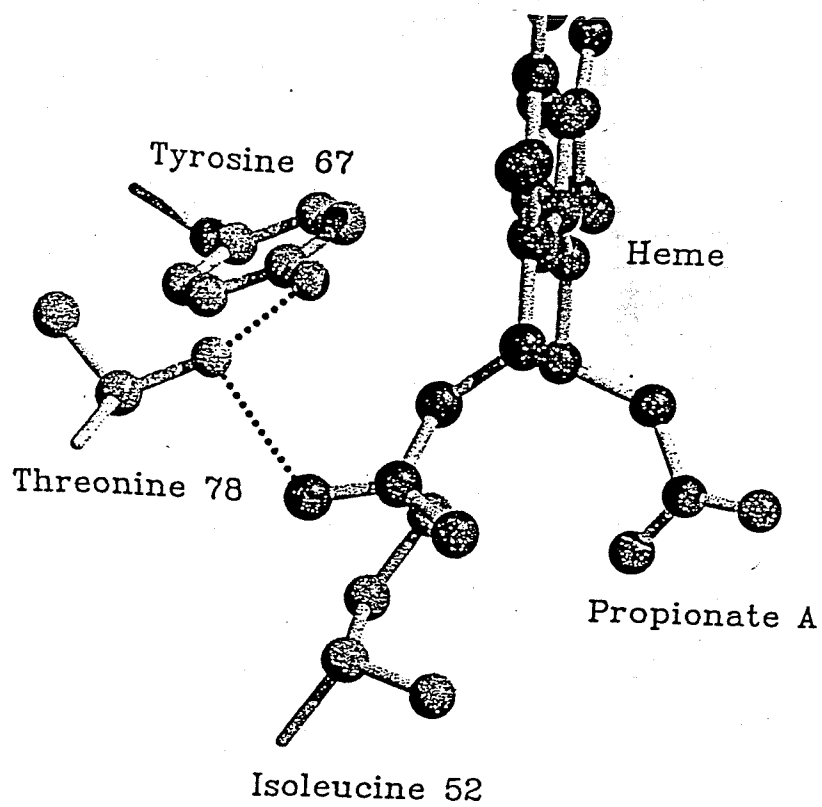
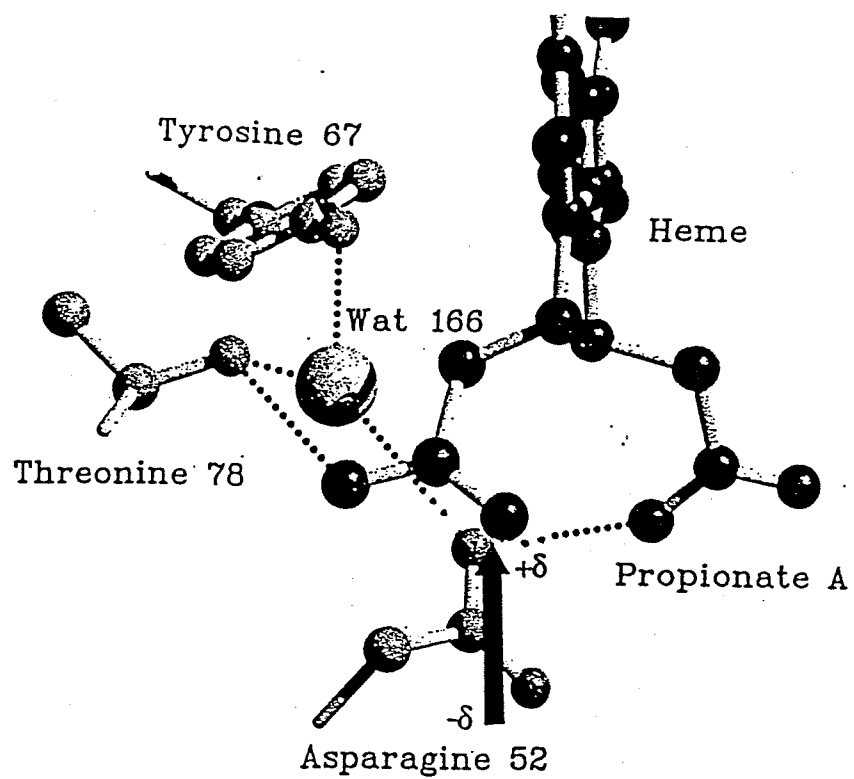


Figure 4.5. Showing the convergence of the average electrostatic energy (ΔG_{elec}) of the native (O) and the mutant protein (U) for trajectories, that were initiated at the corresponding X-ray structures, as well as for the trajectory of the mutant protein, that started from the X-ray structure of the native protein(n). The calculations demonstrated, that the trend in the change in ΔG_{elec} upon mutation may be estimated without having the X-ray structure of the mutant protein.

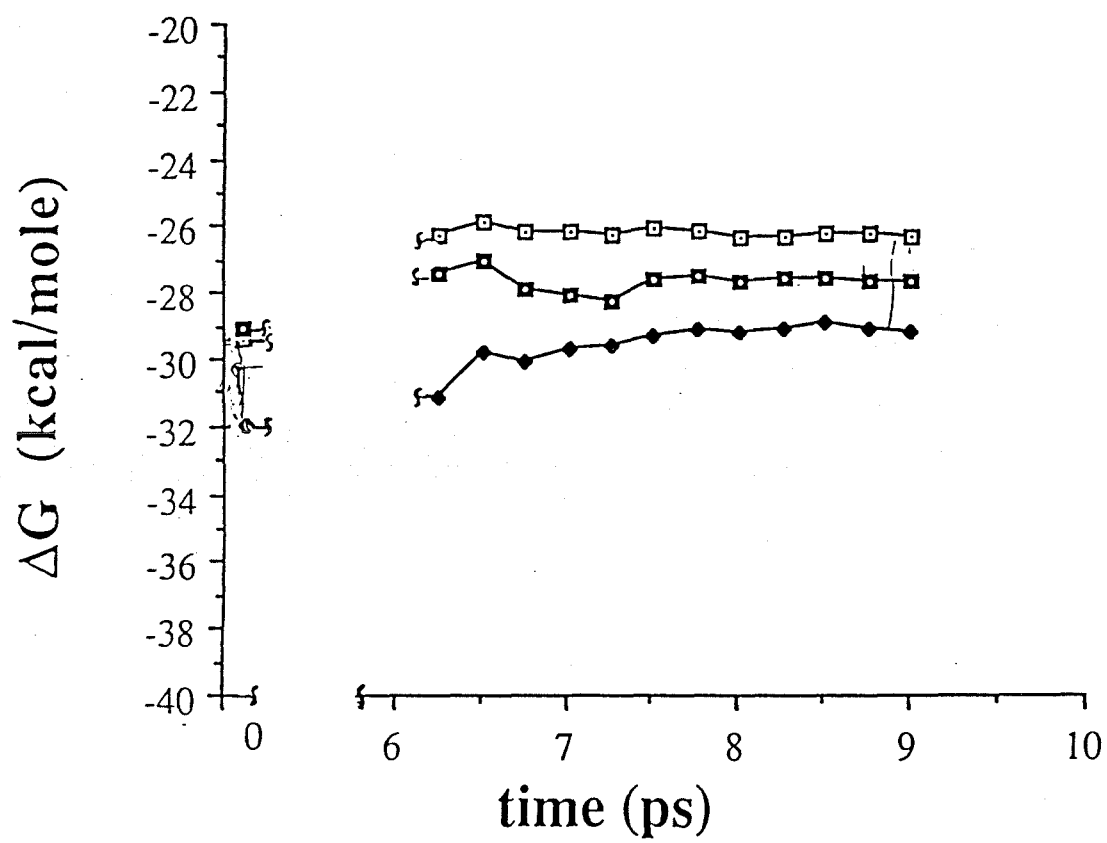
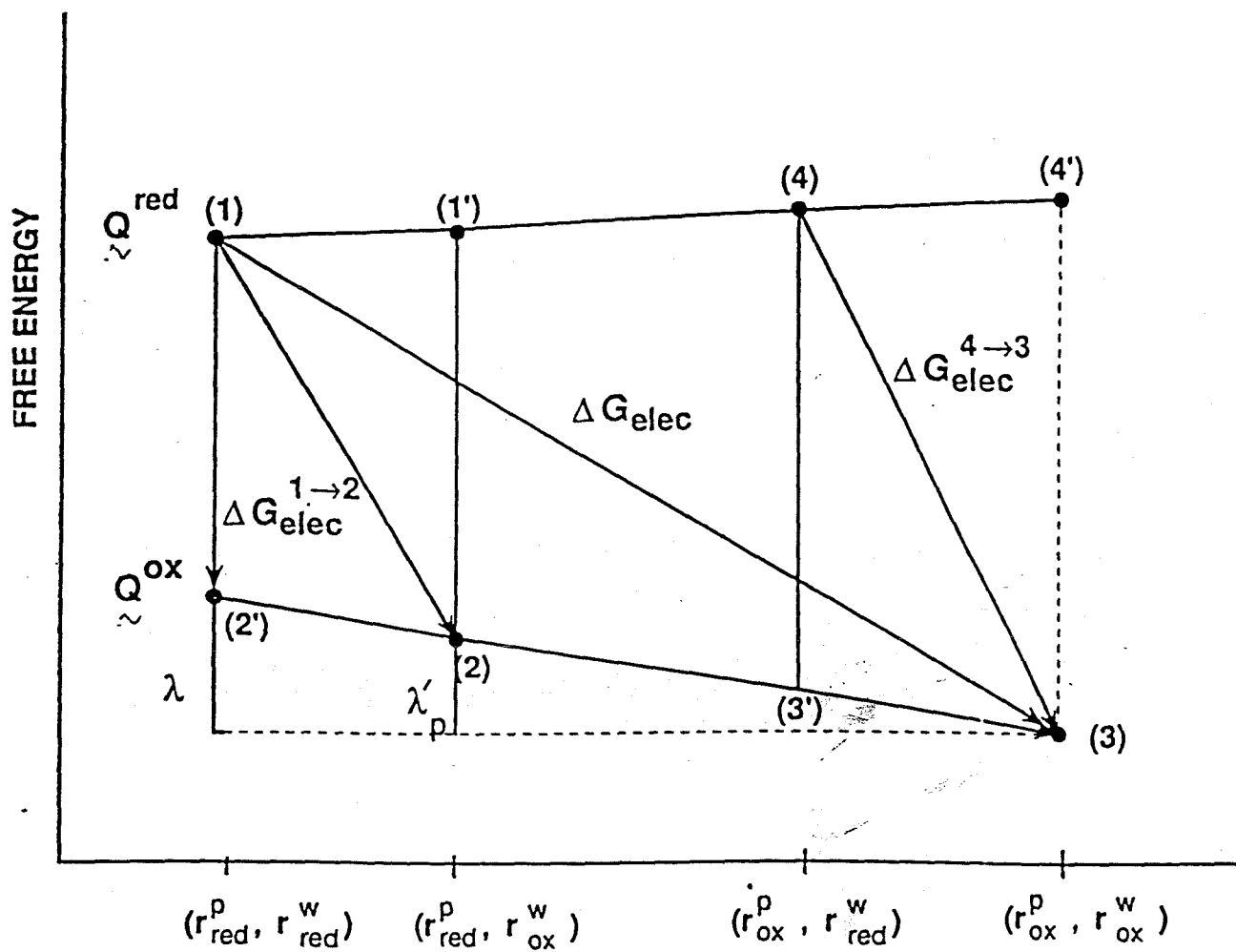


Figure 4.6. The relationship between the energetics of the reduced and oxidized cytochromes and the coordinate of the protein (\mathbf{r}^P) and the solvent (\mathbf{r}^W). The figure does not consider the gas phase energy $\alpha^{(g)}$ and describes the energetics of the charges of the reduced and oxidized hemes (Q^{red} and Q^{ox}) in their cythochrome site. As seen from the figure the total reorganization energy is the change in energy of the oxidized cytochrome upon change from $(\mathbf{r}_{\text{red}}^P, \mathbf{r}_{\text{red}}^W)$ to $(\mathbf{r}_{\text{ox}}^P, \mathbf{r}_{\text{ox}}^W)$ which means a relaxation of both the protein and water coordinates from their equilibrium values at the reduced state to that in the oxidized state. Similarly, λ_p' is the change in energy of the oxidized cytochrome upon change from $(\mathbf{r}_{\text{red}}^P, \mathbf{r}_{\text{ox}}^W)$ to $(\mathbf{r}_{\text{ox}}^P, \mathbf{r}_{\text{ox}}^W)$.



Chapter 5

Simulation of the Change in Reduction Potential for the M80H Mutation in Cytochrome c

ABSTRACT

This chapter addresses the effect of the axial ligand replacement M80H in cytochrome *c*. The mutation results in a significant decrease in the reduction potential from 260 mV to 40 mV. The resulting reduction potential is now very similar to the reduction potential in cytochrome *b* which contains the same axial ligands (His/His). This makes it possible to choose cytochrome *b5* as a reference system. We tested the notion that the axial ligands in both proteins have similar effects on the reduction potential by a computational method. If the effect of the axial ligands is identical in both proteins, the remaining difference in reduction potential should reflect the different environment in the two proteins (after correction for the fact that two different types of heme groups are present in the two proteins). The effect of the protein environment in these proteins was evaluated by the Protein Dipoles Langevin Dipoles (PDL) method in combination with a molecular dynamics (MD) approach. The simulations show that indeed the large change in reduction potential upon the M80H mutation can be almost entirely attributed to the bonding of the axial ligand rather than any electrostatic contribution.

INTRODUCTION

Electron transfer (ET) is a key reaction in photosynthesis and respiration[1-8]. Both processes utilize a series of proteins that make up the electron transport chain. Each member in those chains has very precisely tuned redox potentials. Not only is the control of the reduction potentials very important for efficient energy conversion, but it also has a major impact on ET kinetics as it determines the driving force[9]. How do proteins manage to control the reduction potentials of their cofactors? It has been proposed that in the case of cytochromes redox potential is controlled by three main factors [10]:

(1) the nature of the heme group, (2) the effect of the axial ligand of the heme iron , (3) the electrostatic nature of the environment around the heme group (determining the self-energy).

(1) Different types of hemes (types a, b, or c) have different substituents with different electron withdrawing properties. It has been documented that these substituents can influence the potential of the heme by several tenth of a Volt [11]. (2) Typically, the exchange of an axial ligand can result in bigger effects [12, 13]. Unfortunately, the prediction of these ligand effects *a priori* is still not fully possible. In the case of the blue copper proteins, however, it has been shown that there is a direct correlation between the high reduction potentials and the ligand field strength [14]. (3) Another major contribution to the reduction potential can come from the environment around the heme group. It has been proposed that the high reduction potential of cytochrome *c* can be viewed as a dielectric effect [15]. According to this macroscopic model the low dielectric environment distabilizes the oxidized state. This conclusion was confirmed by microscopic calculations [16]. In the case of a non polar environment the agreement between a macroscopic model and a microscopic model might not be too surprising, as the missing charge stabilization can be expressed in terms of a low dielectric constant. In the opposite case, however, the stabilization of a charged cofactor in a protein the

macroscopic model is not expected to give a correct description. If the solvent is seen as a high dielectric medium, a charge would always be less stable in a medium of a low dielectric constant such as a protein. Thus, it seems that a correct description of this case needs to consider the protein dipoles (for a more detailed discussion see chapter 1 and ref. [17]). An example of such a case will be given in chapter 6, where it will be shown that the reduction potentials of iron-sulfur clusters can vary by as much as a 1 V depending on the protein environment [18]. In that study the effect of the protein was found to be charge-destabilizing in the case of the HiPIP protein, and charge-stabilizing for the ferredoxins.

This chapter examines the effect of the M80H mutation in cytochrome *c*. This axial ligand replacement reduces the redox potential from 260 to 41 mV [12, 13]. Interestingly, the redox potential is now more similar to cytochrome *b5* (5 mV) which has the same axial ligands (His/His). We use a previously developed microscopically model [16, 19-22] to analyze this experiment. The calculations confirm the notion that the change in redox potential is indeed mainly due to a "direct bonding effect" of the axial ligand. In contrast, the change in the electrostatic environment around the heme group (ligand exchange as well as secondary effects) has a negligible contribution to the change in reduction potential. Our study also indicates that the lower redox potential in cytochrome *b5* compared to the H80 mutant is due to the different heme environments in the respective proteins. The heme in cytochrome *b5* is more solvent exposed than in cytochrome *c*. This leads to a more efficient stabilization of the oxidized heme in cytochrome *b5* and therefore to a lower redox potential.

THEORY AND METHOD

As can be seen in Fig. 5.1 the difference in redox potential between the native cytochrome *c* and the mutated cytochrome *c* is 220 mV, whereas the difference between the mutated cytochrome *c* and the cytochrome *b5* is only 35 mV. If the observed change in redox potential upon mutation was totally due to the effect of the axial ligand (which is assumed to be the same in the mutated cytochrome *c* and the native cytochrome *b5*), one would expect the effect of the protein plus water environment around the heme group and the two different heme types to be responsible for the remaining difference in redox potential between the mutated cytochrome *c* and the wild-type cytochrome *b5*. The effect of the different heme types has already been determined experimentally (20 to 30 mV) leaving the environment contribution the only undetermined quantity. If the above assumptions are valid, this environment effect is expected to be 55 to 65 mV (35 mV + 20 to 30 mV). Thus, calculations of the solvation of the heme group by its micro environment (protein plus water) can be directly used to assess the contributions that lead to the observed change in redox potential upon axial ligand replacement.

In order to quantify these solvation effects we used the previously developed protein dipoles Langevin dipoles (PDL) method [16, 19-22]. It evaluates the electrostatic energy for the reduced and oxidized heme group in its protein plus water environment.

The difference in electrostatic (or solvation) energy between the reduced and oxidized ($\Delta\Delta G_{elec}$) can be approximated by

$$\Delta\Delta G_{elec} \approx \Delta V_{Qq} + \Delta V_{Q\alpha} + \Delta V_{Qw} \quad (1)$$

where ΔV_{Qq} , $\Delta V_{Q\alpha}$, and ΔV_{Qw} are the energy changes due to the electrostatic interactions of the heme group (Q) with the permanent charges on the protein atoms (q), with the protein induced dipoles (α) and with the surrounding water (w).

The PDL method was used here in three different ways. The first set of calculations compares ΔG_{elec} for the wild-type structures of cytochrome *c* and cytochrome *b5*. The electrostatic energy contributions from all surrounding water molecules (within 18 Å) and protein residues were evaluated with the exception of the axial ligands in the respective proteins. This simplification seemed appropriate, since the effect of the axial ligands in the two proteins was assumed to be the same and it would therefore not contribute to the redox potential difference.

In another approach we replaced M80 in the wild-type cytochrome *c* by His using the old M80 geometry. This time the electrostatic effects of all residues (and surrounding water molecules within 18 Å) including the axial ligands was evaluated. Note that the effects of bonding to the axial ligand are not included since they are not of electrostatic nature.

The approaches as outlined above were based on the wild-type X-ray structures of the native proteins. However, possible effects of dynamics as well as structural changes upon mutation were not included yet. We therefore checked the calculations by a method that has already successfully been used in previous cases (see also chapters 4) [23]. This technique allowed us to take the structure of the native protein and to simply replace a residue of interest by another residue using the geometry of the replaced residue. The generated structure was then minimized by molecular dynamics (MD). After 5 ps of MD simulations (in the oxidized state) snapshots of the mutant protein were collected for use in PDL calculations. The same MD/PDL calculations were done with cytochrome *b5* to obtain reference points.

RESULTS AND DISCUSSION

Fig. 5.3a shows the direct evaluation of the difference of ΔG_{elec} (in kcal/mol) for the respective wild-type proteins using the PDL method. The effect of the axial ligand is not included in these calculations. The calculated difference is -1.81 kcal/mol, or 78 mV difference in redox potential due to the environment contributions. This becomes roughly 50 mV when the effect of the different heme groups is taken into account. This means that the difference in protein plus water solvation between the two wild-type proteins cytochrome *c* and cytochrome *b* is relatively small compared to the large difference in redox potential between these two proteins. Therefore, it is possible to estimate the difference due to the axial ligand to be approximately 200 mV. Thus, these results confirm that the effect of the different ligands is the main factor in the difference in redox potential between the wild-type cytochrome *c* and the wild-type cytochrome *b*.

A closer look at the individual contributions towards the calculated difference shows that ΔV_{Qq} in cytochrome *c* is about 8.3 kcal/mol more negative than in cytochrome *b*. It is important to note that the relatively large numbers for ΔV_{Qq} are caused by the charges of the heme propionic acids, which would stabilize the heme group significantly in a low dielectric medium. However, as in previous studies, a big compensating effect for both proteins becomes apparent, indicating a relatively high effective dielectric constant inside the proteins. This compensation arises from two contributions, the induced dipoles and the water molecules inside and outside of the protein. Cytochrome *c* shows a large compensation due to the induced dipoles ($\Delta V_{\text{Q}\alpha} = 14.19$ kcal/mol), whereas cytochrome *b* does not. However, both proteins have a large compensation due to the water molecules as the positive ΔV_{bulk} and the positive ΔV_{QL} contributions indicate. This compensation is over 4 kcal/mol larger in cytochrome *b*. Comparing the structures of the two proteins, it appears that cytochrome *c* has a higher compensating effect due to its

induced dipoles, whereas cytochrome *b5* has a higher compensating effect due to its water molecules. In cytochrome *b5* the heme propionic acids are exposed to the solvent water molecules, whereas cytochrome *c* has protein residues surrounding the heme group.

Previous experiments and calculations have shown that an all-water solvent is about 7 kcal/mol better at solvating an oxidized heme group than the protein groups in cytochrome *c* [15, 16]. Thus, it is not surprising that the protein with the more solvent accessible heme group has a higher total ΔG_{elec} .

The second approach is shown in Fig. 5.3b. These calculations are also based on the wild-type X-ray structures for both proteins but the M80 in cytochrome *c* has been replaced by a His residue using the previous M80 geometry. The compensation pattern is very similar to the results from Fig. 5.3a. The difference in ΔG_{elec} of -1.21 kcal/mol is equivalent to 52 mV, which is almost identical with the value expected in the case where the bonding of the axial ligand is responsible for the observed shift upon axial ligand exchange. However, this difference must be viewed as a rather crude estimate, because the mutant structure was not energy minimized.

Structural changes upon mutation might at least partially have contributed to the observed change in redox potential. This effect was included in a third set of calculations. Here we used the X-ray structure of the wild-type protein as the starting point in generating snapshots of mutant structures for PDL calculations. The results obtained by averaging the ΔG 's of the wild-type and the mutant proteins are shown in Figure 5.2. As in previous calculations [23] the average values converge very fast after about 3 to 4 snapshots. And the final ΔG difference of -1.83 kcal/mol (or 79 mV) appears to be quite stable throughout the whole monitored relaxation process. This value is in very close

agreement with the one calculated using the wild-type structures. Thus, the structural changes due to the mutation do not seem to have a strong effect on the energetics of the heme solvation by the protein plus water system.

The calculations which represent a unique way of calibrating ligand effects confirm that the axial ligand itself controls the redox potential in the two cytochromes. About 200 mV of the 255 mV difference between wild-type cytochrome *c* and cytochrome *b5* can be attributed to the different axial ligation. These results suggest that the axial ligand could be mainly responsible for the magnitudes of redox potentials, whereas the protein effect could be in fine-tuning these potentials. However, in the case of iron-sulfur proteins, it is the different protein environment that shift the potentials by more than 1 V (chapter 6)[18].

REFERENCES

1. Warshel, A., *J. Phys. Chem.* **86**, 2218 (1982)
2. Bertini, I., Gray, H.B., Lippard, S.J., and Valentine, J.S.: *Bioinorganic Chemistry*, Mill Valley, CA: University Science Books; 1994.
3. Sigel, H., and Sigel, A.: *Metal Ions in Biological Systems*, New York: Dekker; 1991.
4. Salemme, F.R., *Ann. Rev. Biochemistry* **46**, 299 (1977)
5. Mayo, S.L., Ellis, W.R., Crutchley, R.J., and Gray, H.B., *Science* **233**, 948 (1986)
6. Gray, H.B., and Malmström, B.G., *Biochemistry* **28**, 7499 (1989)
7. Warshel, A., and Parson, W.W., *Ann. Rev. Phys. Chem.* **42**, 279 (1991)
8. McLendon, *Acc. Chem. Res.* **21**, 160 (1988)
9. Marcus, R.A., and Sutin, N., *Biochim Biophys. Acta* **811**, 265 (1985)
10. Marchon, J.C., Mahiko, T., and Reed, C.: *Electron Transport and Oxygen Utilization*, New York: Elsevier North Holland; 1982.
11. Falk, J.E.: *Porphyrins and Metalloporphyrins*, New York: Elsevier; 1964.
12. Rafael, A., and Gray, H.B., *J. Am. Chem. Soc.* **113**, 1038 (1991)
13. Rafael, A., and Gray, H.B., *Proteins* **6**, 338 (1989)
14. Gray, H.B., and Malmström, B.G., *Comments Inorg. Chem.* **2**, 203 (1983)
15. Kassner, R.J., *Proc. Natl. Acad. Sci* **69**, 2263 (1972)
16. Churg, A.K., and Warshel, A., *Biochemistry* **25**, 1675 (1986)

17. Warshel, A., and Aqvist, J., *Ann. Rev. Biophys. Biophys. Chem.* **20**, 267 (1991)
18. Langen, R., Jensen, G.M., Jacob, U., Stevens, P.J., and Warshel, A., *J. Biol. Chem.* **267**, 25625 (1992)
19. Churg, A.K., Weiss, R.M., Warshel, A., and Takano, T., *J. Phys. Chem.* **87**, 1683 (1983)
20. Warshel, A., and Russel, S., *Quart. Rev. Biophys.* **17**, 283 (1984)
21. Russel, S., and Warshel, A., *J. Mol. Biol.* **185**, 389 (1985)
22. Warshel, A., Chu, Z.T., and Parson, W.W., *Science* **246**, 112 (1989)
23. Langen, R., Brayer, G.D., Berghuis, A.M., McLendon, G., Sherman, F., and Warshel, A., *J. Mol. Biol.* **224**, 589 (1992)

Figure 5.1. Redox Potential in mV (vs SHE). Wild-type and mutant are denoted by wt and mut respectively.

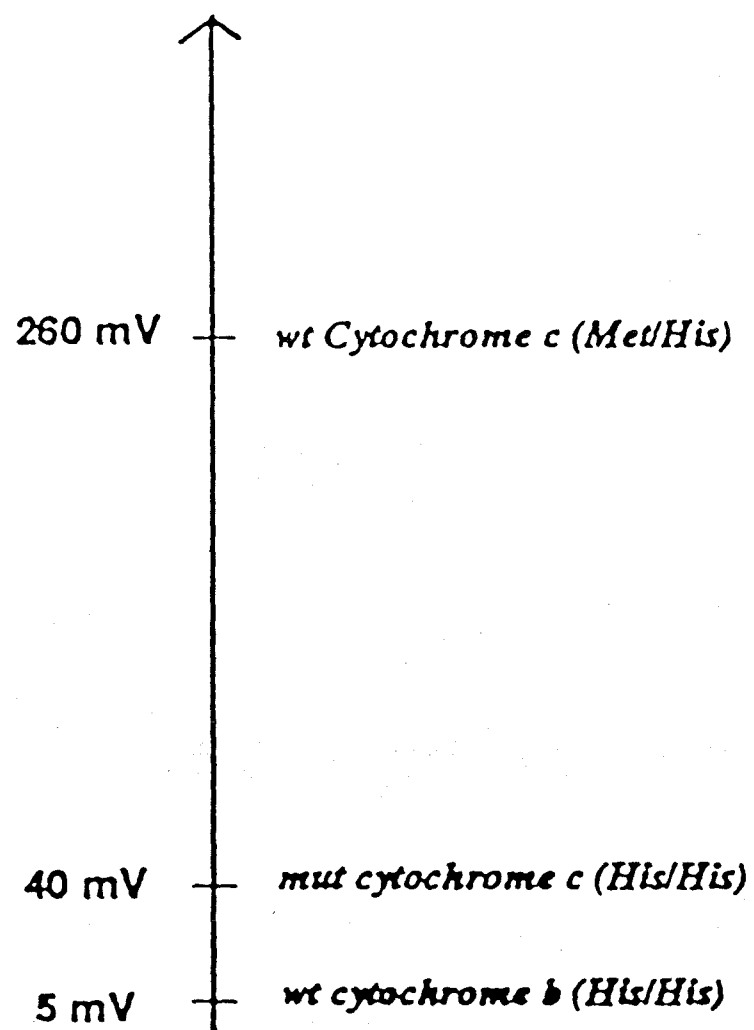


Figure 5.2. Showing the convergence of the averaged electrostatic energies (ΔG_{elec}) of cytochrome *b5* (solid circles) and His-80 cytochrome *c* (open circles). Trajectories were initiated at the corresponding X-ray structure of the wild-type proteins and run in the oxidized state. In the case of the His-80 cytc mutant the X-ray structure was modified by simply replacing the Met-80 coordinates with His.

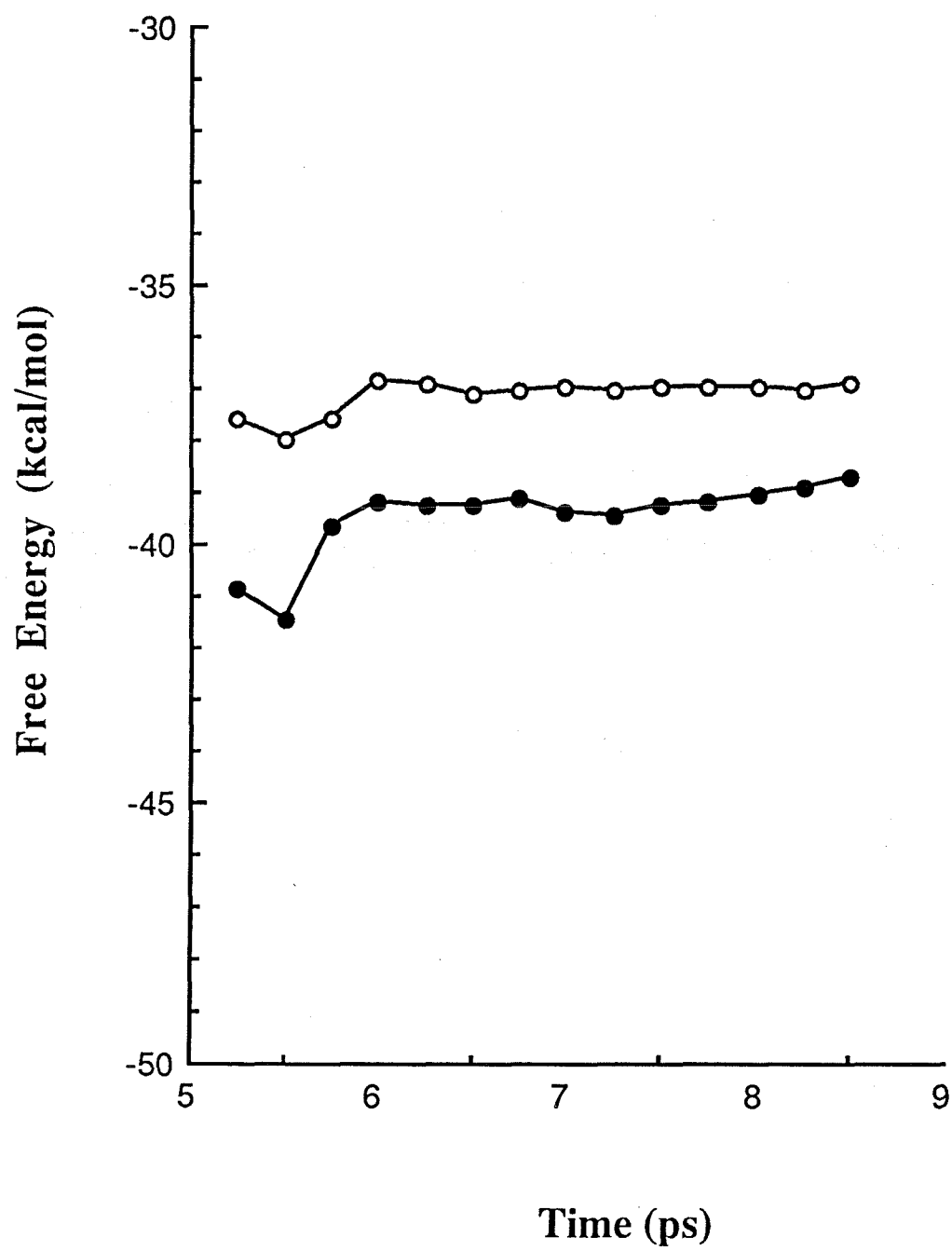


Figure 5.3. Comparison of the solvation energy contributions for the oxidation of cytochrome *c* and cytochrome *b*. The individual terms are described in the methodology sections of chapters 4 and 5.

Figure 5.3. Comparison of the solvation energy contributions for the oxidation of cytochrome *c* and cytochrome *b*. The individual terms are described in the methodology sections of chapters 4 and 5.

	ΔV_{Qq}	$\Delta V_{Q\alpha}$	ΔG_{QL}	ΔG_{bulk}	ΔG_{elec}	$\Delta\Delta G$
--	-----------------	----------------------	-----------------	--------------------------	--------------------------	------------------

a) Axial ligands uncharged

Cytochrome c (H80)	-84.06	14.19	0.75	36.68	-32.43	-1.81
Cytochrome b	-75.66	-0.07	4.98	36.51	-34.24	

b) Axial ligands charged

Cytochrome c (H80)	-83.49	14.59	0.80	36.68	-31.41	-1.21
Cytochrome b	-78.18	1.87	7.18	36.51	-32.62	

Chapter 6

Protein Control of Iron Sulfur Cluster Redox Potentials

[This work is described in Langen, R., Jensen, G.M., Jacob, U., Stevens, P.J., and Warshel, A. (1992) *J. Biol. Chem.*, **267**, 25625-25627]

ABSTRACT

Iron-sulfur proteins constitute a major class of redox active metalloproteins. The relationship between the three-dimensional structures of iron-sulfur proteins and their corresponding redox potentials is of fundamental importance. We report successful calculations of the redox potentials of the $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]^{-2/-3}$ couple in several crystallographically characterized proteins using a microscopic electrostatic model which includes both protein and water. Our results confirm the important role of amide groups close to the cluster, a suggestion made previously. However, proteins with extremely similar amide group arrangements still exhibit a wide range of redox potentials. Our results show that in these cases proteins can tune the redox potential of this couple in part by varying the access of solvent water to the cluster. The simple electrostatic model employed provides the first successful quantitative modelling of the protein control of iron-sulfur redox potentials.

INTRODUCTION

Proteins containing iron-sulfur ([Fe-S]) clusters occur ubiquitously in nature and play a major role as biological electron carriers [1]. Thus, the understanding of the control of the redox potential of an [Fe-S] cluster by its protein environment is of fundamental importance. This chapter addresses the origin of the enormous variation in the midpoint potential of the $-2/-3$ couple of $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]$ clusters of four small, structurally well-characterized proteins: Azotobacter vinelandii ferredoxin I (AvFdI), Peptococcus aerogenes ferredoxin (PaFd), Bacillus thermoproteolyticus ferredoxin (BtFd) and Chromatium vinosum high potential iron protein (CvHiPIP). The electrostatic interaction of the [Fe-S] clusters with their protein and water environment is calculated using the "protein dipoles Langevin dipoles" (PDL) approach [2-4]. The predicted variation in the electrostatic contribution of protein and water together to the cluster redox potential is in agreement with experiment. These calculations constitute the first successful modelling of protein control of [Fe-S] cluster redox potentials.

The structures of AvFdI, PaFd, BtFd and CvHiPIP have been determined by X-ray crystallography [5-11]. AvFdI contains one $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]^{-2/-3}$ cluster and one $[\text{Fe}_3\text{S}_4(\text{S-cys})_3]^{-2/-3}$ cluster whose midpoint potentials are $\sim -650\text{mV}$ and $\sim -420\text{mV}$ (vs SHE) respectively [12]. PaFd contains two $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]^{-2/-3}$ clusters whose midpoint potentials are both $\sim -430\text{mV}$ [13]. BtFd contains a single $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]^{-2/-3}$ cluster; the midpoint potential of the nearly identical Bacillus stearothermophilus ferredoxin is -280mV [14]. CvHiPIP contains a $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]$ cluster whose $-2/-3$ couple has not yet been expressed under physiological conditions; instead, the $-1/-2$ couple is active [15, 16]. The correlation of the structures and redox potentials of these proteins presents a major challenge.

In this work we evaluate the effect of the protein on the redox potential of $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]$ clusters in the above proteins using the PDL approach [2-4] as implemented in

the program POLARIS [17]. The electrostatic interaction of a charged moiety - in this case, the $[\text{Fe}_4\text{S}_4(\text{S}_\gamma)_4]^{-2/-3}$ ($\text{S}_\gamma \equiv$ cysteine S) portion of the [Fe-S] cluster - with its surroundings is expressed as the sum of four terms: $V_{Q\mu}$, $V_{Q\alpha}$, V_L and V_B . $V_{Q\mu}$ and $V_{Q\alpha}$ are the interactions of the cluster with the partial charges and induced dipoles of protein atoms. V_L and V_B are interactions with solvent water, treated microscopically (as "Langevin dipoles") within a sphere of radius r_L , centered at the cluster, and macroscopically beyond. Langevin water dipoles are placed on a grid and oriented self-consistently in the combined field from cluster charges, protein charges and induced dipoles and other water dipoles.

Calculations have been carried out for the -2/-3 couple of the $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]$ clusters of AvFdI, BtFd and CvHiPIP and of one of the two clusters of PaFd (that ligated by cysteines 8, 11, 14 and 46); an error in the protein sequence used in the PaFd structure determination in the neighborhood of the second cluster has recently come to light [18] and we have excluded this cluster from consideration. The differential solvation energies of the -2 and -3 clusters obtained are given in Table 1 (which also includes further details of the calculations) and Figure 6.1. Redox potentials obtained thence are compared to experiment in Figure 6.2. The absolute redox potential can be evaluated using the cluster in solution as a reference [4]. Here, however, we only address the trend within the different proteins and we arbitrarily set the calculated and experimental redox potentials of AvFdI to be identical. We predict an ordering of potentials: CvHiPIP << AvFdI < PaFd < BtFd, in agreement with experiment. The spread for AvFdI, PaFd and BtFd is 447 mV compared to an experimental range of 370 mV.

The contributions of $V_{Q\mu}$, $V_{Q\alpha}$, V_L and V_B to the total differential solvation energies of BtFd, PaFd, AvFdI and CvHiPIP are given in Table 1 and are diagrammed in Figure 6.1. The much greater contribution of $V_{Q\mu}$ in BtFd, PaFd and AvFdI than in CvHiPIP is the primary cause of the much higher potential of the -2/-3 couple of the former three

proteins. The crystallographic studies of BtFd, PaFd, AvFdI and CvHiPIP have identified hydrogen (H)-bonding by amide NH groups to the $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]$ clusters. The first three proteins exhibit highly conserved cluster environments with H-bonding from 8 amide groups [5,8,18,19]. In the case of CvHiPIP 5 amide groups are involved in H-bonding [10,18]. Elimination of the charges of these amide groups leads to the results given in Table 1. The total differential solvation energies are substantially reduced in BtFd, PaFd and AvFdI; in contrast, in CvHiPIP very little change occurs. These results demonstrate the importance of both the presence and the orientation of the amide groups in the immediate environment of the $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]$ clusters of BtFd, PaFd and AvFdI in elevating their redox potentials above that of CvHiPIP. The difference in cluster-amide H-bonding between CvHiPIP and PaFd was recognized early as a likely major contributor to their very different redox properties [19, 20]. The results confirm, refine and quantitate this expectation.

Although the contribution of the local amide groups accounts for the difference between CvHiPIP and PaFd it does not explain the variation in redox potential within the BtFd, PaFd and AvFdI group. This variation originates in changes in the contributions of $V_{Q\mu}$, $V_{Q\alpha}$ and V_L which are comparable in magnitude. The combined contributions of $V_{Q\mu}$ and $V_{Q\alpha}$ are PaFd < AvFdI ~ BtFd while for V_L AvFdI < BtFd < PaFd. Thus, the lower potential of AvFdI compared to PaFd is attributable to a substantially smaller V_L contribution which outweighs the greater $V_{Q\mu}$ and $V_{Q\alpha}$ contribution. In contrast, the lower potential of PaFd relative to BtFd is attributable to the substantially lower $V_{Q\mu}$ and $V_{Q\alpha}$ contribution which outweighs the greater V_L contribution. The inter-relationship of the variations in the protein structure and in $V_{Q\mu}$, $V_{Q\alpha}$ and V_L contributions to the redox potential is complex. Calculations in which protein groups more than 9Å from the cluster center are discarded exhibit very similar variations, showing that variations in redox potential originate predominantly in changes in protein structure in the neighborhood of the cluster. Of particularly significant interest is the fact that the water contribution V_L is

the dominant factor in causing the difference in redox potential between AvFdI and PaFd. The environments of the clusters in these two proteins exhibit strong homology [5, 18]. Apparently, however, the extent of water penetration to the neighborhood of the clusters is significantly different, as illustrated in Fig. 6.3.

The results provide further support for the PDL approach in describing electrostatic energetics in aqueous solutions of proteins. Note that in the calculations (i) there is no change in structure with oxidation state, (ii) all residues (excepting cysteine cluster ligands) are uncharged and (iii) the second cluster in PaFd and AvFdI are uncharged. The inclusion of protein reorganization is expected to improve the agreement of calculation and experiment (see, e.g., ref. 21). The effects of including cluster interactions with ionized residues (and other clusters) are expected to be relatively small. Estimates are given in Table 1.

Although the factors governing the redox potentials of [Fe-S] clusters have been widely discussed [1, 18], microscopic calculations of protein tuning of [Fe-S] cluster redox potentials have not been reported previous to this work. Furthermore, the crucial role of the solvent in establishing the trend in redox potentials has not been demonstrated. In view of the agreement between our calculations and experiment, we are extending our studies to encompass other [Fe-S] clusters/redox couples [1] and to examine the results of site-specific mutations in AvFdI [12, 22, 23].

REFERENCES

1. H. Beinert, *FASEB J.* 4, 2483 (1990).
2. A. Warshel and S.T. Russell, *Quart. Rev. Biophys.* 17, 283 (1984).
3. A. Warshel and J. Åquist, *Ann. Rev. Biophys. Biophys. Chem.* 20, 267 (1991).
4. A.K. Churg and A. Warshel, *Biochem.* 25, 1675 (1986).
5. C.D. Stout, *J. Mol. Biol.* 205, 545 (1989).
6. E.T. Adman, L.C. Sieker and L.H. Jensen, *J. Biol. Chem.* 251, 3801 (1976).
7. K. Fukuyama, Y. Nagahara, T. Tsukihara and Y. Katsube, *J. Mol. Biol.* 199, 183 (1988);
8. K. Fukuyama, H. Matsubara, T. Tsukihara and Y. Katsube, *J. Mol. Biol.* 210, 383 (1989).
9. C.W. Carter Jr. *et al.*, *J. Biol. Chem.* 249, 4212 (1974)
10. C.W. Carter Jr., J. Kraut, S.T. Freer and R.A. Alden, *J. Biol. Chem.* 249, 6339 (1974).
11. S.T. Freer, R.A. Alden, C.W. Carter Jr., and J. Kraut, *J. Biol. Chem.* 250, 46 (1975).
12. S.E. Iismaa *et al.*, *J. Biol. Chem.* 266, 21563 (1991).
13. N.A. Stombaugh, J.E. Sundquist, R.H. Burris and W.H. Orme-Johnson, *Biochem.* 15, 2633 (1976).
14. R.N. Mullinger *et al.*, *Biochem J.* 151, 75 (1975).
15. C.W. Carter Jr. *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 69, 3526 (1972).
16. R. Cammack, *Biochem. Biophys. Res. Comm.* 54, 548 (1973).
17. A. Warshel and S. Creighton in *Computer Simulation of Biomolecular Systems*, W.F. Van Gunsteren and P.K. Weiner, Eds. Leiden (1989) pp. 120-138.
18. G. Backes *et al.*, *J. Am. Chem. Soc.* 113, 2055 (1991).
19. E.T. Adman, K.D. Watenpaugh, L.H. Jensen, *Proc. Natl. Acad. Sci. U.S.A.* 72, 4854 (1975).
20. C.W. Carter Jr. in *Iron-Sulfur Proteins*, W. Lovenberg, Ed. New York (1977) pp. 157-204.
21. R. Langen *et al.*, *J. Mol. Biol.* , 224, 589
22. A.E. Martin *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 87, 598 (1990).

23. J. Soman, S. Iismaa and C.D. Stout, *J. Biol. Chem.* 266, 21558 (1991).
24. L. Baciou, I. Sinning and P. Sebban, *Biochem.* 30, 9111 (1991).
25. G. Lamm and G.R. Pack, *Proc Natl. Acad. Sci. U.S.A.* 87, 9033 (1990).
26. L. Noodleman, J.G. Norman, J.H. Osborne, A. Aizman and D.A. Case, *J. Am. Chem. Soc.* 107, 3418 (1985).

Figure 6.1. Contributions to the differential solvation energies of the $[\text{Fe}_4\text{S}_4(\text{S-cys})_4]^{-2/3}$ clusters of BtFd, PaFd, AvFdI, and CyHiPIP.

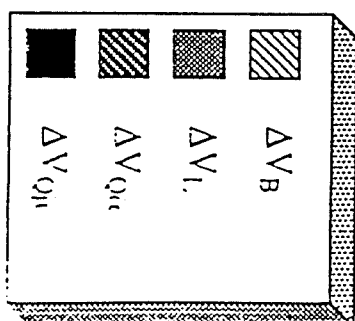
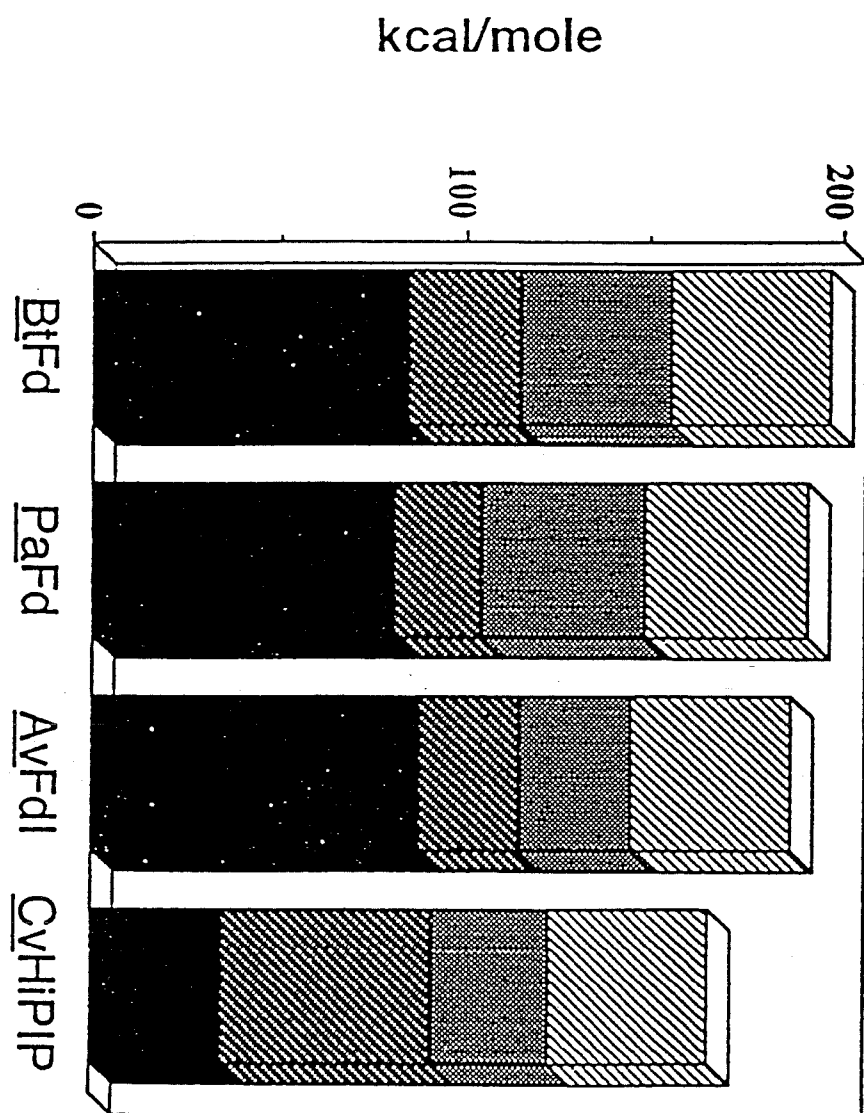


Figure 6.2. Calculated and experimental redox potentials of BtFd, PaFd, AvFdI, and CyHiPIP.

Redox Potential (mV vs. SHE)

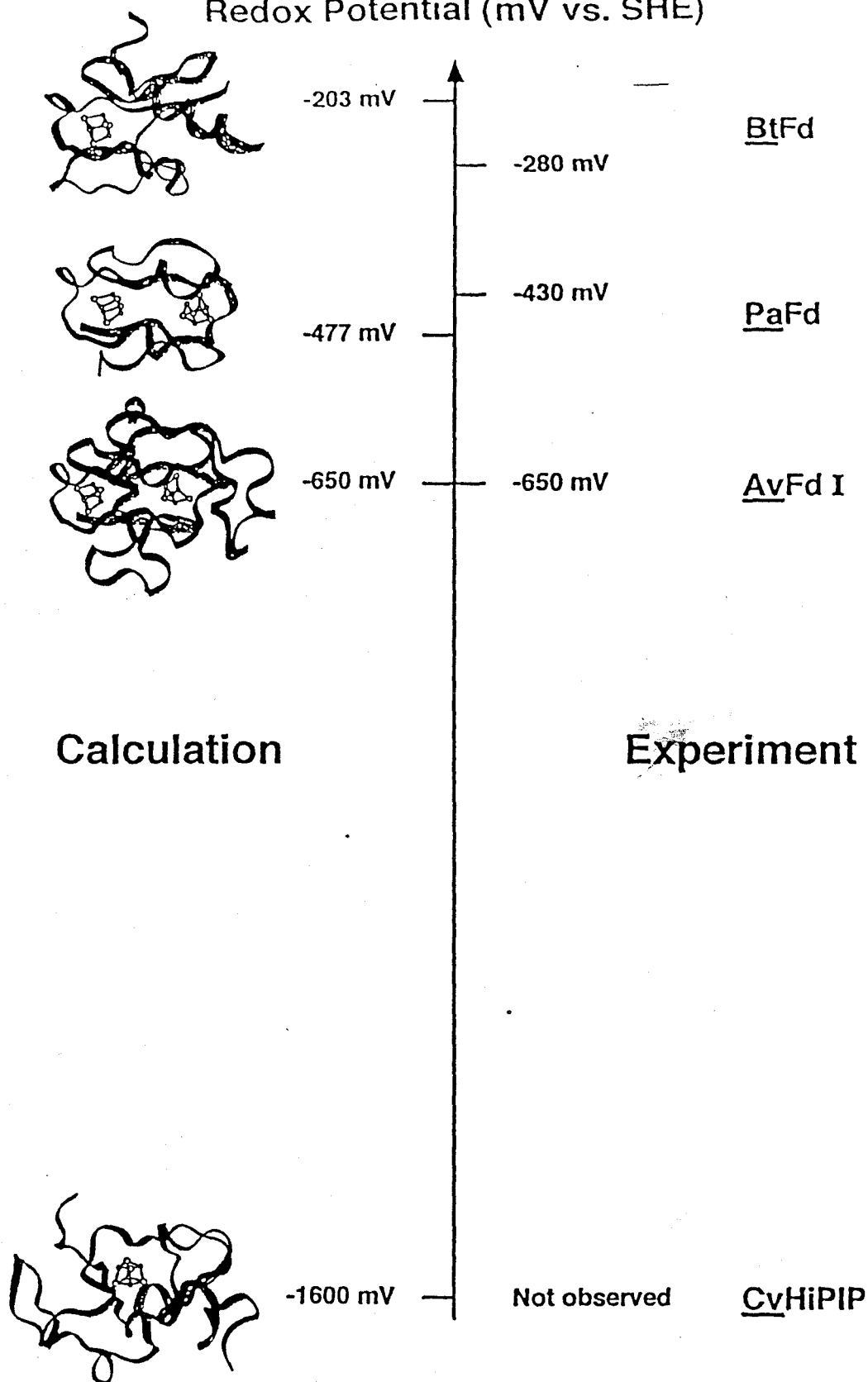
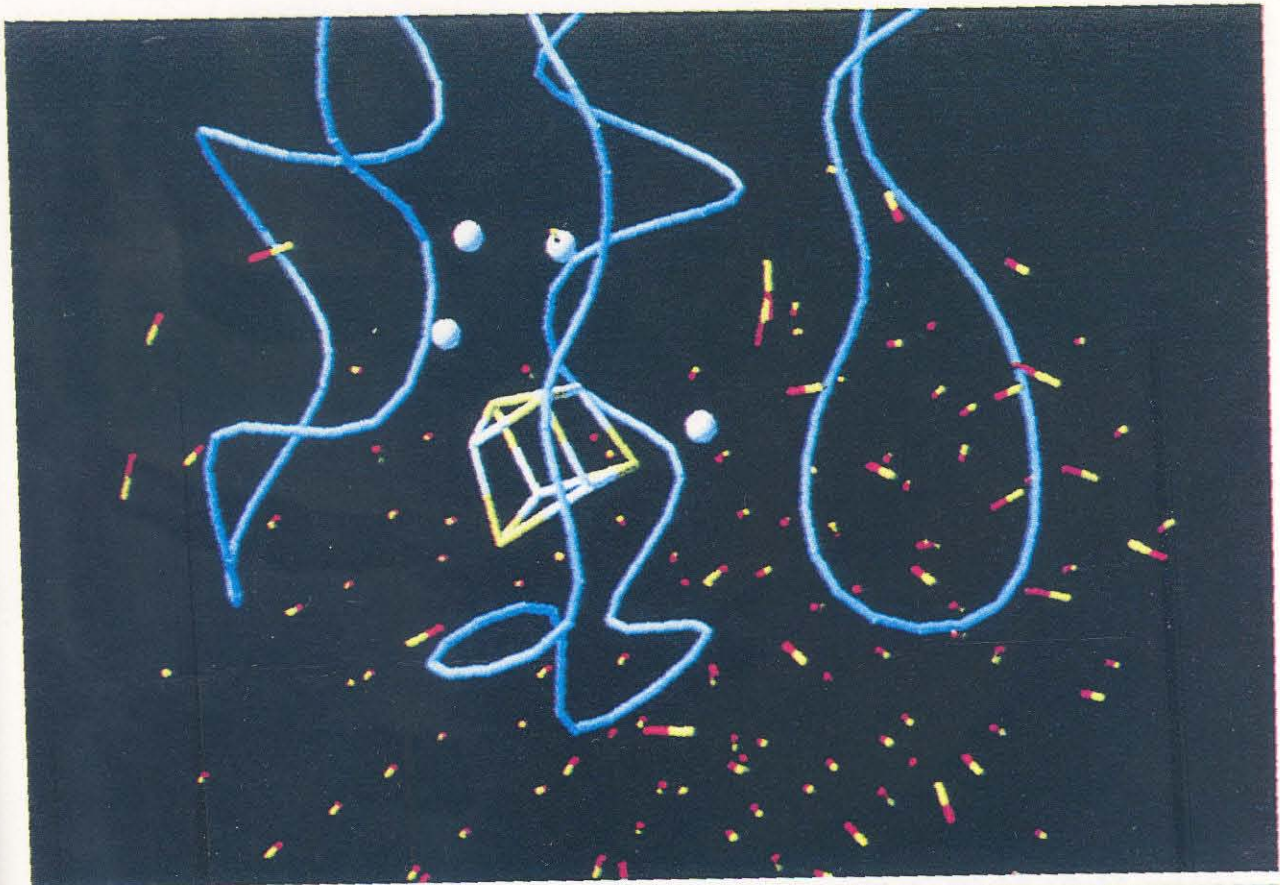


Figure 6.3. The protein/Langevin dipole system for PaFd (upper) and AvFdI (lower). The Langevin dipoles within 6 Å from the clusters are shown as spheres, the rest are depicted as red/yellow dipoles. The main chain regions which are homologous in both proteins are shown in blue, and the additional C-terminal portion of AvFdI is shown in grey. Four dipoles are found within 6 Å of the cluster in PaFd; however, there is no space for Langevin dipoles within 6 Å of the cluster in AvFdI.



Appendix

*Calculations on the GTP-Hydrolysis
Mechanism in Oncoprotein Ras P21*

This appendix presents theoretical studies on the GTP-hydrolysis mechanism in Ras p21. The work uses computational techniques [1-4] similar to the ones presented in Chapters 4-6, however, it is otherwise not related to the main part of this thesis. Ras p21 proteins constitute a group of highly conserved proteins that play a very important role in signalling of cell growth and differentiation in eukaryotes. Ras proteins have the ability to bind GTP and to hydrolyze it to GDP. The interaction with the GTPase activating protein (GAP) can cause a rate enhancement of several orders of magnitude for this reaction [5]. The GTP bound conformation represents the active, growth promoting state of the protein whereas the GDP bound form is biologically inactive.

Mutated (oncogenic) Ras proteins have been found in 90 percent of human pancreatic cancers, 50 percent of colorectal tumors, 30 percent of lung cancers, and 10 to 40 percent of numerous other cancers, making *ras* the most prevalent oncogene in human cancer. The mutations found so far involve only three residues, Gly12, Gly13 and Gln61. As has been shown by X-ray crystallography [6-10] these residues are in close proximity to the γ -phosphate. Biochemically they impair the proteins ability to hydrolyze GTP and render it insensitive to activation by GAP [11]. Therefore, the protein remains for a longer time in an active, growth promoting GTP bound state, which leads to uncontrolled cell growth. Thus, an understanding of the GTPase mechanism and the interaction with GAP appears to be of utmost importance for an understanding of the molecular basis of *ras* related cancer.

A possible mechanism for GTP hydrolysis in p21 has been proposed by Pai et al.[8]. It is based on crystallographic data in which a possibly nucleophilic water molecule was identified. The hydrolysis reaction has been proposed to involve two steps. The first step represents the activation of a water molecule by transferring a proton to Gln 61. In the second step the generated OH^- attacks the γ -phosphate (see Figure 1 in Reprint1).

In Reprint 1 the energetics of the proposed mechanism are evaluated using a theoretical method[1-4]. The calculations suggest that the proposed hydrolysis mechanism is energetically unfavorable, as the protein does not provide enough stabilization for the protonated Q61. The calculations also show that the stabilization of the protonated Q61 by the nearby negatively charged E63 could have been substantial, however the local environment in the protein acts as a high dielectric medium diminishing the interaction from 60 to less than 2 kcal/mol.

In Reprint 2 we reconfirm that a mechanism in which Q61 acts as general base is very unlikely, and we address the question what the actual mechanism might be. Several possibilities were considered. However, only one appeared to be promising. This mechanism involved the protonation of the γ -phosphate with subsequent attack of this group by the generated OH^- .

REFERENCES

1. Warshel, A., Sussman, F., & Hwang, J.-K., *J. Mol. Biol.* **201**, 139-159 (1988)
2. Warshel, A.: *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, New York: John Wiley & Sons; 1991.
3. Åqvist, J., & Warshel, A., *Biochemistry* **28**, 4680-4688 (1989)
4. Åqvist, J., and Warshel, A., *J. Mol. Biol.* **224**, 7 (1992)
5. Bollag, G., and McCormick, F., *Nature* **351**, 576 (1991)
6. De Vos, A.M., Tong, L., Milburn, M., Matias, P.M., Jancarik, J., Noguchi, S., Nishimura, K., Miura, K., Ohtsuka, E., & Kim, S.H., *Science* **239**, 888-893 (1988)
7. Pai, E.F., Kabsch, W., Krengel, U., Holmes, K., John, J., & Wittinghofer, A., *Nature* **341**, 209-214 (1989)
8. Pai, E.F., Krengel, U., Petsko, G.A., Goody, R.S., Kabsch, W. & Wittinghofer, A., *EMBO J.* **9**, 2351-2359 (1990)
9. Prive, G.G., Milburn, M.V., Tong, L., De Vos, A.M., Yamaizumi, Z., Nishimura, S., & Kim, S.H., *Proc. Natl. Acad. Sci* **89**, 3649 (1992)
10. Krengel, U., Schlichting, I., Scherer, A., Schumann, R., Frech, M., John, J., Pai, E.F., & Wittinghofer, *Cell* **62**, 539-548 (1990)
11. Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M., and Wigler, M., *Proc. Natl. Acad. Sci.* **81**, 4008 (1984)

REPRINT I

**On the Mechanism of Guanosine Triphosphate
Hydrolysis in Ras p21 Proteins.**

Langen, R., Schweins, T., and Warshel, A.

Biochemistry (1992) , **31**, 8691-8696

On the Mechanism of Guanosine Triphosphate Hydrolysis in *ras* p21 Proteins[†]

Ralf Langen, Thomas Schweins, and Arieh Warshel*

Department of Chemistry, University of Southern California, Los Angeles, California 90089-1062

Received June 11, 1992; Revised Manuscript Received July 15, 1992

ABSTRACT: The residue Gln61 is assumed to play a major role in the mechanism of *ras* p21, and mutations of this residue are often found in human tumors. Such mutations lead to a major reduction in the rate of GTP hydrolysis by the complex of *ras* p21 and the GTPase activating protein (GAP) and lock the protein in a growth-promoting state. This work examines the role of Gln61 in *ras* p21 by using computer simulation approaches to correlate the structure and energetics of this system. Free energy perturbation calculations and simpler electrostatic considerations demonstrate that Gln61 is unlikely to serve as the general base in the intrinsic GAP-independent reaction of p21. Glutamine is already a very weak base in water, and surprisingly the GlnH⁺ OH⁻ reaction intermediate is even less stable in the protein active site than in the corresponding reaction in water. The electrostatic field of Glu63, which could in principle stabilize the protonated Gln61, is found to be largely shielded by the surrounding solvent. However, it is still possible that Gln61 is a general base in the GAP/*ras* p21 complex since this system could enhance the electrostatic effect of Glu63. It is also possible that the γ -phosphate acts as general base and that Gln61 accelerates the reaction by stabilizing the OH⁻ nucleophile. If such a mechanism is operative, then GAP may enhance the effect of Gln61 by preorienting its hydrogen bonds in the transition-state configuration.

ras p21 proteins constitute a group of highly conserved proteins that play a major role in signaling cell growth and differentiation in eukaryotes. Like other G-proteins these proteins have the ability to bind GTP and hydrolyze it to GDP. Their intrinsic GTPase activity can be enhanced over 1000-fold (Bollag & McCormick, 1991) by physical interaction with the GTPase activating protein (GAP). The GTP-to-GDP conversion leads to a significant conformational change in several regions of the *ras* protein. The GTP-bound conformation seems to represent the active, growth-promoting state of the proteins whereas the GDP-bound form appears to be inactive.

Mutations in positions 12, 13, and 61 of *ras* p21 have been found in a large number of human tumors. X-ray crystallography studies (De Vos et al., 1988; Pai et al., 1989; Krengel et al., 1990; Milburn et al., 1990) have located these residues in close proximity to the GTP γ -phosphate (Figure 1). Biochemical analysis has shown that these mutations impair the protein's ability to hydrolyze GTP (Fasano et al., 1984) and render it insensitive to activation by GAP (Adari et al., 1988). Therefore, the protein remains in the active, growth-promoting GTP-bound state, which can lead to cancer. Thus an understanding of the GTPase mechanism of p21 and the effect of GAP on this mechanism is of importance for an understanding of the molecular basis of cancer. The role of Gln61 in the mechanism of p21 is of particular interest since mutations of this residue in the GAP/*ras* p21 complex reduce the rate of GTP hydrolysis by more than 4 orders of magnitude (Sigal et al., 1988) and are found to be responsible for cell transformation (Fasano et al., 1984). The same mutations in

the isolated p21 change the hydrolysis rate by only 1 order of magnitude (Fasano et al., 1984).

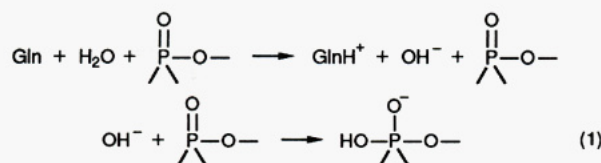
A recent study (Pai et al., 1990; Krengel et al., 1990) proposed that Gln61 helps to facilitate the nucleophilic attack on the γ -phosphate by activating a water molecule. Although the proposed mechanism is not explicitly stated, it seems to imply that Gln61 is the general base in the hydrolysis. Such a mechanism [which is referred to here as the general base (GB61) mechanism] is consistent with the observation of a water molecule between Gln61 and the GTP perfectly positioned for a direct in-line attack on the γ -phosphate. However, despite the appealing structural evidence this mechanism is not fully established. For example, structural studies of p21 mutants have been recently used as an argument against the role of Gln61 as a general base (Prive et al., 1992). In order to validate the proposed mechanism, it is essential to correlate the structure of the system with its energetics.

This work examines the GB61 mechanism by using the structure of *ras* p21 to evaluate the energetics of the assumed proton-transfer process. This is done by using the empirical valence bond (EVB) approach that has been used previously in semiquantitative studies of the closely related hydrolytic reaction of staphylococcal nuclease. The EVB results are further verified by well-defined electrostatic considerations. The present study indicates that Gln61 is not likely to be the general base in the reaction of isolated p21. Alternative mechanisms and the possible effect of GAP are considered as well.

SIMULATION STUDIES

Simulations of the Proton-Transfer Step. The GB61 mechanism can be formally described as

[†] This study was supported by the Gottlieb Daimler and Karl Benz Foundation (to T.S.) and the Charles Heidelberger Memorial Fellowship (to R.L.).



The energetics of this or any other assumed mechanism can be examined by the EVB method [see, e.g., Warshel et al. (1988) and Warshel (1991)]. This method describes the reacting system as a mixture of resonance structures representing different feasible bonding and charge configurations. The potential surfaces of these resonance structures are calibrated by experimental information of solution reactions and then transferred without any change in parameters to the protein environment. The free energy surface for the enzyme reaction is then evaluated by a combined free energy perturbation (FEP) and umbrella sampling formulation. The EVB/FEP approach is described in detail elsewhere (Warshel et al., 1988), and its reliability has been demonstrated in several test cases [see, e.g., Warshel et al. (1991)], including studies of the catalytic reaction of staphylococcal nuclease (Åqvist & Warshel, 1989), which is related to the GB61 mechanism. The parameters and conditions used in the present simulation are given in Table I.

The EVB/FEP free energy surface of the GB61 mechanism for the GTP hydrolysis in *ras* p21 is depicted in Figure 2. Details of the simulations are outlined in the caption of Table II. The overall rate constant of the reaction is determined by the activation barrier for the proton transfer and the nucleophilic attack stages, which are designated by Δg_1^* and Δg_2^* , respectively. Of particular interest is the fact that the overall activation barrier cannot be lower than Δg_1^* . Keeping this in mind, we find it very significant that the calculated value of Δg_1^* is $\sim 30 \pm 4$ kcal/mol, while the observed activation free energy, Δg_{obs}^* , for the overall reaction of p21 is 23 kcal/mol. This value is deduced from the observed rate constant of $3.3 \times 10^{-4} \text{ s}^{-1}$ (Temeles et al., 1985), by using transition-state theory with a preexponential factor of $6 \times 10^{12} \text{ s}^{-1}$, which is found to be applicable for reactions in condensed phases with a significant activation barrier [see, e.g., Warshel et al. (1988), Kraut (1988), and Warshel (1991)]. One may still argue that tunneling correction should reduce Δg_1^* , but such corrections are expected to be similar in the reference reaction and in the protein (Hwang et al., 1991) and Δg_1^* is larger than $\Delta G_{\text{PT}}^{\text{p}}$ even with tunneling correction. This is important, since $\Delta G_{\text{PT}}^{\text{p}}$ is already higher than the measured energy for the whole reaction. Furthermore, the activation energy for the second step of the reaction is larger than zero so that the total activation barrier must be higher than $\Delta G_{\text{PT}}^{\text{p}}$. Thus, the EVB calculations suggest that Gln61 is not the general base, at least in the absence of GAP.

Electrostatic Validation of the Calculated Activation Barrier. The validity of the conclusions of the previous section depends, of course, on the accuracy of our calculations. In the present case it is possible to check the calculations by rather simple but quite reliable electrostatic considerations. One can exploit the fact that not only the activation free energy, Δg_1^* , but even the free energy of the proton-transfer process, $\Delta G_{\text{PT}}^{\text{p}}$, is predicted to be larger than the experimentally determined activation free energy. An estimate of this free energy, which is significantly easier to calculate than Δg_1^* , can be obtained by using the protein dipoles langevin dipoles (PDL) method. This method, described in detail in Warshel and Russell (1984) and Langen et al. (1992), allows one to determine the differences between the electrostatic stabi-

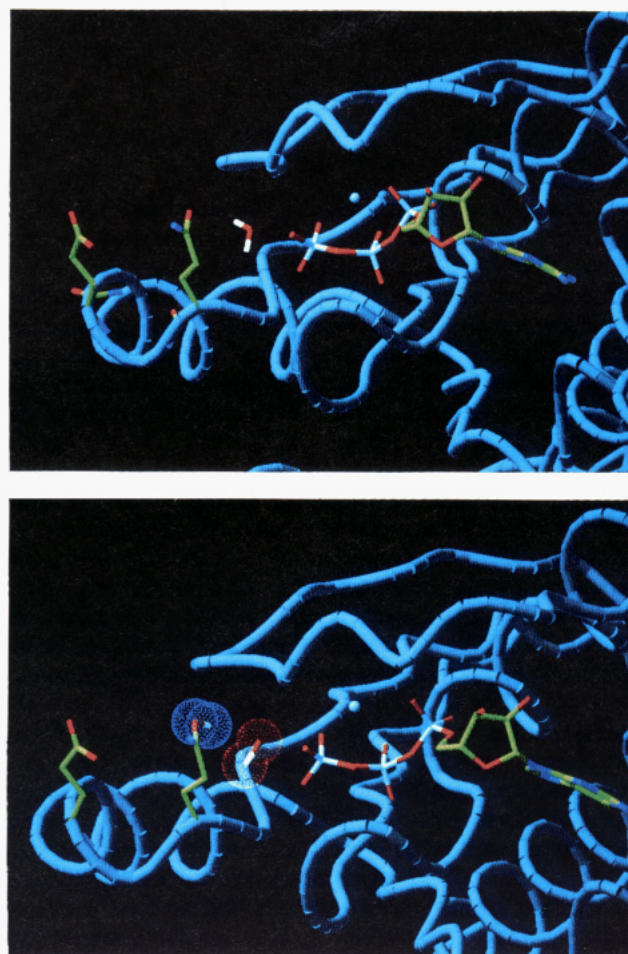


FIGURE 1: (a, top) Structure showing the active site of *ras* p21. Highlighted residues are (from left to right) Glu63, Gln61, water, and GTP. Mg^{2+} is shown as a sphere. (b, bottom) Structure showing the reaction intermediate after the proton-transfer step. Highlighted residues are the same as in (a). The red- and blue-dotted surfaces depict the product of the charge of the given ion and the time-average electrostatic potential from the medium (red indicates positive and blue indicates negative energy contribution). The actual potential energies (relative to the reference water system) are -8 and 17 kcal/mol for the sites of GlnH^+ and OH^- . Thus OH^- is less stable in the protein than in the reference solvent cage while the GlnH^+ is more stable, but this effect is smaller than the destabilization of the OH^- .

zation of the reacting groups in the protein active site and in a water reference system. The corresponding differences in electrostatic energies can then be used in conjunction with the thermodynamic cycle of Figure 3 to determine the free energy of the proton-transfer step. That is, using Figure 3 we can write [see, e.g., Warshel (1981, 1991)]

$$\begin{aligned}
 \Delta G_{\text{PT}}^{\text{p}} &= \Delta G_{\text{PT}}^{\text{w}} + \Delta G_{\text{sol},2}^{\text{w} \rightarrow \text{p}} - \Delta G_{\text{sol},1}^{\text{w} \rightarrow \text{p}} \\
 &= \Delta G_{\text{PT}}^{\text{w}} + \Delta \Delta G_{\text{sol}}^{\text{w} \rightarrow \text{p}} \quad (2)
 \end{aligned}$$

where $\Delta G_{\text{sol}}^{\text{w} \rightarrow \text{p}}$ designates the change in the corresponding "solvation energy" or electrostatic energy upon moving the reacting species in a given configuration from water to the protein active site. To evaluate eq 2, it is necessary to know the free energy of the proton-transfer step in solution. This free energy can be obtained to a very good approximation from the difference between the pK_a 's of the donor and acceptor groups [e.g., Warshel (1981, 1991)], using

$$\begin{aligned}
 \Delta G_{\text{PT}}^{\text{w}} &\simeq -2.3RT[\text{pK}_a(\text{GlnH}^+) - \text{pK}_a(\text{H}_2\text{O})] \\
 &\simeq 23 \text{ kcal/mol} \quad (3)
 \end{aligned}$$

This energy is already as large as Δg_{obs}^* , reflecting the fact

Table I: Parameters Used in the Calculations^a

Charges: ^b $V_{qq} = 332q_iq_j/r_{ij}$			
(HO-C-NH ₂) ⁺ (ψ_2, ψ_3)	$q_C = 0.301, q_O = -0.101, q_N = -0.202,$ $q_{H1} = 0.366, q_{H2} = q_{H3} = 0.318$		
(-O-POO ₂) ²⁻ (ψ_1, ψ_2)	$q_{O1} = -0.48, q_{O2} = q_{O3} = q_{O4} = -0.82,$ $q_P = 0.94$		
(-O(OH)-POO ₂) ²⁻ (ψ_3)	$q_{O1} = q_{O2} = -0.5, q_{O3} = q_{O4} = q_{O5} = -1.0,$ $q_P = 1.0$		
Nonbonded: ^b $V_{nb} = A_iA_j/r^{12} - B_iB_j/r^6$			
	$A_{Mg^{2+}} = 96$	$B_{Mg^{2+}} = 32$	
Charged Residues ^c			
Asp38, 33, 57	$q_{Asp} = -1$	Mg	$q_{Mg} = +2$
Lys16, 117, 88	$q_{Lys} = +1$	GTP (ψ_1, ψ_2)	$q_{GTP} = -4$
Arg68	$q_{Arg} = +1$	GTP-OH (ψ_3)	$q_{GTP-OH} = -5$
Glu62, 63	$q_{Glu} = -1$		

^a Energies are in kcal/mol, distances in Å, and atomic charges in atomic units. Parameters not listed are the same as in Warshel et al. (1988) and Åqvist and Warshel (1989). ^b The sensitivity of the final results to the parameters used was checked by repeating the calculations with modified parameters (e.g., changing the residual charges by 15%). The effect on the final results was minor since similar changes occurred in the energies of the reaction in the protein and in solution. ^c The effect of distant ionized groups was estimated using the macroscopic Coulomb law and a high dielectric constant (Warshel & Russell, 1984). The corresponding effect on the final result is rather small.

that Gln is a weak base in water. Thus, if the free energy for proton transfer in the protein, ΔG_{PT}^w , is similar to ΔG_{PT}^w , then Δg_1^* in the protein is larger than Δg_{obs}^* . Of course, the protein groups can change ΔG_{PT} ; this protein effect is given by $\Delta\Delta G_{sol}^{w \rightarrow p}$ of eq 2, which is the change in electrostatic energy of the reacting system upon moving from water to the protein active site. In particular, if the protein microenvironment destabilizes the GlnH⁺ OH⁻ ion pair (relative to water), then $\Delta\Delta G_{sol}^{w \rightarrow p}$ is positive and ΔG_{PT}^p is larger than ΔG_{PT}^w . The PDL calculations (Table II) indicate that the protein destabilizes the ion pair and that ΔG_{PT}^p is ~25 kcal/mol, which again is larger than the experimentally determined activation free energy for the whole reaction.

The present results might appear somewhat unexpected, given the fact that the nearby Glu63 could in principle greatly stabilize the protonated form of Gln61. This issue can also be examined by the PDL calculations, as summarized in Figure 4. As seen from the figure, the overall effect of the negatively charged Glu63 on the protonated Gln61 is only ~2 kcal/mol. The reason for this unexpectedly small stabilization is the remarkable compensation between charge-charge interactions and solvation effects (Warshel & Russell, 1984; Warshel & Åqvist, 1991). That is, while the "gas-phase" interaction between Glu63 and the protonated Gln61 is very large (~65 kcal/mol), this interaction is compensated in the protein by an almost equal contribution (~63 kcal/mol) from the protein permanent and induced dipoles and the surrounding water molecules. In other words, the field of Glu63 is shielded quite effectively by the surrounding medium. Such an effect, which is equivalent to a local, high dielectric constant, is quite common for surface groups (Warshel & Åqvist, 1991) and has been analyzed in detail in related cases [see, e.g., Churg and Warshel (1986)].

The conclusion that the proton transfer to Gln61 is not facilitated by the protein can also be deduced by considering the relative electrostatic potentials of the GlnH⁺ OH⁻ ion pair in solution and in p21. These potentials, described in Figure 1 in terms of the corresponding electrostatic energies, demonstrate that the OH⁻ ion is destabilized by *ras* p21 relative to water. Apparently, the GlnH⁺ is stabilized by the protein, but this effect is smaller than the destabilization of the OH⁻

ion, so that overall the ion pair is less stable in the protein. The relative destabilization of the OH⁻ ion is due to the negative charges of the γ -phosphate whose field is shielded in water more effectively than in the protein. This is true despite the fact that the Mg²⁺ ion and other positively charged residues (e.g., Lys16) decrease the negative potential of the γ -phosphate in *ras* p21.

DISCUSSION

The present simulation studies indicate that Gln61 is unlikely to be the base in the intrinsic reaction of p21¹. This finding is consistent with the experimental observation of Der et al. (1986), who showed that all mutations at this position lead to an 8–10-fold reduction of the reaction rate. These are quite small changes, as compared to the effects of mutations in proteins where a residue is directly involved in the catalytic process [see, e.g., Wilkinson et al. (1984) and Sepersu et al. (1987)]. Such mutations usually lead to a reduction of the reaction rate by several orders of magnitude as in the case of staphylococcal nuclease, where the corresponding effect (Sepersu et al., 1987) was quantitatively reproduced by EVB calculations (Åqvist & Warshel, 1989). Furthermore, mutation of Gln to residues which are far better bases, e.g., Glu or His, would have been expected to accelerate the reaction, but such an effect has not been observed.

The X-ray structures of Milburn et al. (1990) and Pai et al. (1989) are not identical, and our simulation started from the coordinates of Milburn et al. (1990). Thus one may wonder about the generality of the present results. Our FEP approach uses the X-ray structure only as its starting point and then explores many configurations, including those of the transition-state region which are not identical to the X-ray structure. Similarly, our approach can generate the position of the water molecules rather than using those observed in the X-ray structure (usually we find water molecules at the X-ray positions). With this procedure we obtain similar energies, to within 4 kcal/mol, for simulations with significantly different initial coordinates. Thus we believe that similar results would have been obtained with the coordinates of Pai et al. (1989).

The mechanism proposed by Pai et al. (1990) emphasizes the possible role of the main-chain carbonyl of Thr35 which is hydrogen bonded to the nucleophilic water molecule. Our calculations include this group, as well as any other group of the protein, and found that its catalytic role is not so obvious. Although the main-chain carbonyl group of Thr35 polarizes the catalytic water, it does not help the overall proton-transfer process, since it destabilizes the generated OH⁻ ion. Nevertheless, the carbonyl dipole of Thr35 can reorient itself and stabilize the protonated Gln61. The resulting stabilization is small because the energy gain through this movement is partly compensated by reorganization energy effects, an effect frequently encountered in group reorientation.

If Gln61 is not the general base for the intrinsic GTPase activity of p21, then we are left with only few other proton acceptors. In fact, the X-ray structure does not show an obvious residue in an appropriate position, and no candidate base has yet been identified by site-specific mutagenesis.

¹ After completing the main part of this work (Schweins, 1991), we became aware of a related proposal (Prive et al., 1992) that interprets structural studies of mutations at position 61 and other considerations as evidence against the GB61 mechanism. However, this insightful work did not use energy considerations to correlate the structural information with the energetics of the GB61 mechanism.

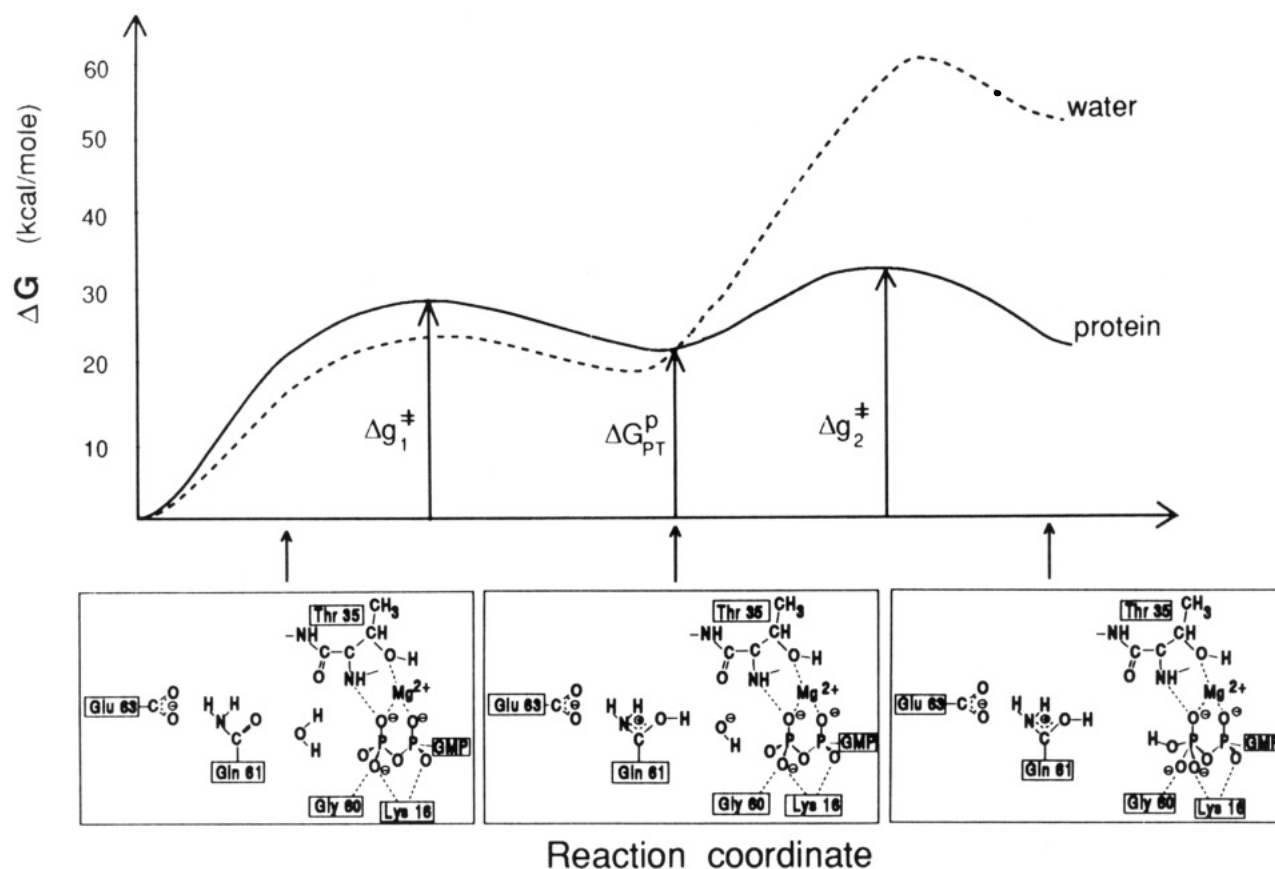


FIGURE 2: Calculated EVB/FEP reaction profiles of the GB61 mechanism in p21 and in a reference solvent cage. Δg_1^\ddagger and ΔG_{PT}^p are respectively the activation barrier and the free energy for proton transfer from a water molecule to Gln61. Superscripts p and w designate protein and water. The lower part of the figure schematically describes representative configurations along the reaction coordinate.

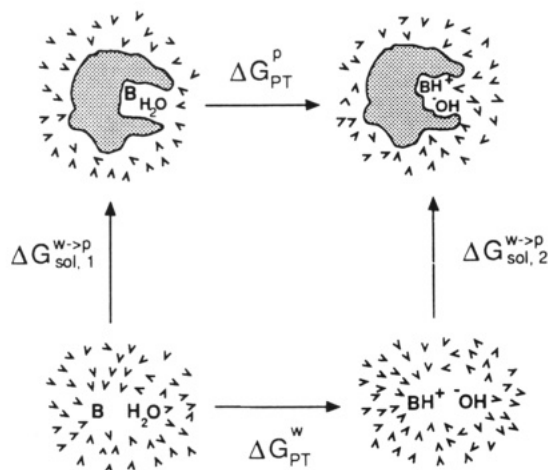


FIGURE 3: Thermodynamic cycle relating the free energies of the proton-transfer step in solution to the corresponding energetics in the protein's active site.

Another nearby water molecule could act as an alternative base. Of course, nucleophilic attack might also precede the proton-transfer step, or both steps may be concerted. It is possible that Asp57 is a general base for a water molecule that is bound to the Mg^{2+} ion. However, the nucleophilic attack of this water seems to be inconsistent with the observation of in-line attack and inversion of the γ -phosphate center (Feuerstein et al., 1989).

The γ -phosphate itself might also serve as a proton acceptor since the phosphate anion could be a stronger base than water. This depends, however, on the effect of the Mg^{2+} ion. Gln61 could play a major role in this mechanism by stabilizing the generated negative charge rather than by serving as base. For

Table II: Energetics of the Proton-Transfer Step from Water to Gln61^a

method	process	energy in protein		energy in water	
		ΔG_{PT}^p	$(\Delta g_1^\ddagger)^p$	ΔG_{PT}^w	$(\Delta g_1^\ddagger)^w$
EVB/FEP	$B + H_2O \rightarrow BH^+ + OH^-$	25	30 ± 4	22	26 ± 3
PDL	$B + H_2O \rightarrow BH^+ + OH^-$	23		22	

^a Energies in kcal/mol. The EVB/FEP simulations were carried out as described by Warshel et al. (1989) and with the recent modifications of Lee et al. (1992), using the ENZYME program (Warshel & Creighton, 1989). The simulation involves 11 mapping steps of 2 ps each with 2-fs time steps at 300 K. Cutoff radii of 15, 20, and 18 Å were used for the water sphere, the explicit protein region, and the surrounding grid of Langevin dipoles [see also Lee et al. (1992)]. The PDL calculations were done with the POLARIS program (Warshel & Creighton, 1989), averaging the calculations over eight protein configurations. The estimated error range reflects the average results of several simulations and the uncertainty in the exact activation barrier in solution.

example, the amide group can provide hydrogen bond stabilization to the OH^- nucleophile. In fact, as was shown in the related studies of serine proteases [see, e.g., Hwang and Warshel (1987) and Warshel et al. (1988)], such a stabilization can reduce the activation barrier by as much as 5 kcal/mol. This maximum contribution can be obtained only if the hydrogen bond donor is already prealigned in the optimal direction toward the negative charge (Hwang & Warshel, 1987; Warshel et al., 1988).

A recent work (Prive et al., 1992) has proposed that Gln61 could, in principle, contribute to catalysis by stabilizing the γ -phosphate in the transition-state conformation. However, one has to keep in mind that the distance between the amide group and the γ -phosphate is about 6 Å. Moving Gln61 from

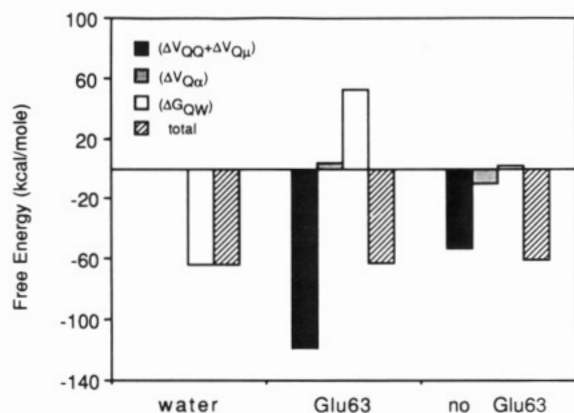


FIGURE 4: Effects of Glu63 on the energy contributions involved in the interaction between OH^- and Gln61. ΔV_{QQ} is the Coulombic interaction between GlnH⁺ and OH^- . $\Delta V_{Q\mu}$ is the interaction of these groups with all the other protein charges. $\Delta V_{Q\alpha}$ is the interaction of GlnH⁺ and OH^- with induced dipoles in the protein. ΔG_{QW} is the interaction with the solvent. This figure illustrates how the very large charge-charge interaction between GlnH⁺ and Glu63 is almost completely compensated for by the change in solvation energy, such that the total energy remains almost constant.

its equilibrium position toward the γ -phosphate may cost a significant amount of energy. The contribution of this residue in stabilizing the pentacoordinated phosphate transition state might be less than the corresponding contribution of nearby water molecules, which would stabilize the transition state when Gln61 is kept at its original place. Here again, one has to consider the concept of compensation through reorganization energy (Warshel, 1991). A significant catalytic effect is hard to obtain when the relevant residue is not at the correct position and orientation in the ground state. The examination of the effectiveness of such a mechanism in isolated p21 will be checked by the EVB/FEP approach.

One of the reasons why the protein is unable to accelerate the GB61 reaction relative to the water reference system appears to be the large compensation for the Glu63–Gln61 interaction. This compensation is mainly due to surrounding water molecules, which shield the field of Glu63 very effectively. Thus, a replacement of these solvent molecules by a medium with a lower local dielectric constant may decrease the free energy and the activation free energy of the proton-transfer reaction. It is possible that the binding of GAP to *ras* involves removal of the water molecules around Glu63 and creation of an environment that leads to a stronger electrostatic interaction between Glu63 and Gln61. Therefore, for the GAP-induced GTPase reaction a mechanism with Gln61 as a base cannot be ruled out. In fact, such a model could explain why all mutations at position 61 make it impossible for *ras* to be activated by GAP, while the effect of these mutations on the intrinsic GAP-independent GTPase activity is relatively small. Such a model may also explain why mutations at position 63 lead to transforming activity (Fasano et al., 1984). Unfortunately, we still do not have any kinetic data on the hydrolysis with and without GAP for these mutants.

Another way by which GAP could help p21 to catalyze the GTP hydrolysis is by controlling the orientation of Gln61 in the protein. As the X-ray studies reveal (Pai et al., 1990; Krenkel et al., 1990), both Gln61 and Glu63 are located in a flexible loop which has a higher thermal mobility. GAP could “freeze” this loop in a catalytic conformation so that Gln61 becomes oriented in a way that requires a small reorganization for stabilization of the transition state (probably by interacting with the OH^- nucleophile).

An instructive experiment that should be very informative is the examination of the actual effect of mutation of Glu63 on the GTP hydrolysis reaction of p21, in the absence of GAP. The effect of Glu63 on the GB61 mechanism can be estimated quite accurately; since this group is well solvated, its effect can be evaluated in a reliable way by using Coulomb's law with a dielectric constant between 40 and 80 (Warshel & Russell, 1984; Warshel & Åqvist, 1991) as well as by the PDL method. Both approaches give a stabilization effect of approximately 1–2 kcal/mol, which corresponds to 1 order of magnitude change in the rate constant. Therefore, although the predicted change in energy is rather small, the corresponding change in rate constant should be reproducible in a mutation experiment. Thus, mutation of Glu63 should change the rate by more than a factor of two if the GB61 mechanism is operative. If such effect is not detected, then the GB61 mechanism can be excluded.

While elucidation of the actual mechanism of p21 and GAP/p21 may still take some time, we have demonstrated here how one can use computer modeling approaches to examine an assumed mechanism. A combination of structural studies, site-specific mutagenesis, and computer modeling approaches may help in the establishment of the catalytic reaction of p21. A good example is the above prediction of the effect of Glu63 in the GB61 mechanism.

ACKNOWLEDGMENT

We are grateful to Prof. S. H. Kim for the coordinates of p21 which were obtained by his group (with the support of NIH Grant CA45593). T.S. gratefully acknowledges support from the Gottlieb Daimler and Karl Benz Foundation.

REFERENCES

- Adari, H., Lowy, D. R., Der, C. J., & McCormick, F. (1988) *Science* 240, 518–521.
- Åqvist, J., & Warshel, A. (1989) *Biochemistry* 28, 4680–4688.
- Bollag, G., & McCormick, F. (1991) *Nature* 351, 576–579.
- Chung, A. K., & Warshel, A. (1986) *Biochemistry* 25, 1675–1681.
- Der, C., Finkel, T., & Cooper, G. (1986) *Cell* 44, 167–176.
- De Vos, A. M., Tong, L., Milburn, M., Matias, P. M., Jancarik, J., Noguchi, S., Nishimura, K., Miura, K., Ohtsuka, E., & Kim, S. H. (1988) *Science* 239, 888–893.
- Fasano, O., Aldrich, T., Tamanoi, F., Taparowsky, E., Furth, M., & Wigler, M. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81, 4008–4012.
- Feuerstein, J., Goody, R. S., & Webb, M. (1989) *J. Biol. Chem.* 264, 6188–6190.
- Hwang, J.-K., & Warshel, A. (1987) *Biochemistry* 26, 2669–2673.
- Hwang, J.-K., Chu, Z.-T., Yadav, A., & Warshel, A. (1991) *J. Phys. Chem.* 95, 8445–8448.
- Kraut, J. (1988) *Science* 242, 533–540.
- Krenkel, U., Schlichting, I., Scherer, A., Schumann, R., Frech, M., John, J., Pai, E. F., & Wittinghofer, A. (1990) *Cell* 62, 539–548.
- Langen, R., Brayer, G. D., Berghuis, A. M., McLendon, G., Sherman, F., & Warshel, A. (1992) *J. Mol. Biol.* 224, 589–600.
- Lee, F. S., Chu, Z. T., Bolger, M. B., & Warshel, A. (1992) *Protein Eng.* 5, 215–228.
- Milburn, M., Tong, L., De Vos, A. M., Bruenger, A., Yamaizumi, Z., Nishimura, S., & Kim, S. H. (1990) *Science* 247, 939–945.
- Pai, E. F., Kabsch, W., Krenkel, U., Holmes, K., John, J., & Wittinghofer, A. (1989) *Nature* 341, 209–214.
- Pai, E. F., Krenkel, U., Petsko, G. A., Goody, R. S., Kabsch, W., & Wittinghofer, A. (1990) *EMBO J.* 9, 2351–2359.

- Prive, G. G., Milburn, M. V., Tong, L., De Vos, A. M., Yamaizumi, Z., Nishimura, S., & Kim, S. H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 3649–3653.
- Schweins, T. (1991) Diplomarbeit Universität Hannover.
- Sepersu, E. H., Shortle, D., & Mildvan, A. S. (1987) *Biochemistry* 26, 1289–1300.
- Sigal, I. S., Marshall, M. S., Schaber, M. D., Vogel, U. S., Scolnick, E. M., & Gibbs, J. B. (1988) *Cold Spring Harbor Symp. Quant. Biol.* 53, 863–869.
- Temeles, G. L., Gibbs, J. B., D'Alonzo, J. S., Sigal, I. S., & Scolnick, E. M. (1985) *Nature* 313, 700–703.
- Warshel, A. (1981) *Biochemistry* 20, 3167–3177.
- Warshel, A. (1991) *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, John Wiley & Sons, New York.
- Warshel, A., & Russell, S. (1984) *Q. Rev. Biophys.* 17, 283–422.
- Warshel, A., & Creighton, S. (1989) in *Computer Simulations of Biomolecular Systems* (Van Gunsteren, W. F., & Weiner, P. K., Eds.) p 120, ESCOM, Leiden, The Netherlands.
- Warshel, A., & Åqvist, J. (1991) *Annu. Rev. Biophys. Biophys. Chem.* 20, 267–298.
- Warshel, A., Sussman, F., & Hwang, J.-K. (1988) *J. Mol. Biol.* 201, 139–159.
- Wells, J. A., Cunningham, B. C., Graycar, T. P., & Estelle, D. A. (1986) *Philos. Trans. R. Soc. London, A* 317, 415–423.
- Wilkinson, A. J., Fersht, A. R., Blow, D. M., Carter, C., & Winter, G. (1984) *Nature* 307, 187–188.
- Registry No.** GTP, 86-01-1; Gln, 56-85-9; Glu, 56-86-0; GTPase, 9059-32-9; H₂O, 7732-18-5.

REPRINT II

**Why have Mutagenesis Studies not Located
the General Base in Ras p21?**

Schweins, T, Langen, R., and Warshel, A.

Nature Structural Biology (1994) , 1, 476-484

Why have mutagenesis studies not located the general base in ras p21

T. Schweins, R. Langen and A. Warshel

Ras p21 plays a major role in the control of cell growth, and oncogenic mutations of this protein have been found in human cancers. Unfortunately, the detailed mode of action of Ras p21 is still unclear, in spite of the great interest in this protein and the availability of its X-ray crystal structure. In particular, mutagenesis studies of different active site residues could not identify the general base for GTP hydrolysis. Here we tackle this question using a computer simulation approach with clear and reliable energy considerations and conclude that the most likely general base is the bound GTP itself. Obviously, the identification of such a general base cannot be easily accomplished by mutagenesis experiments.

Ras p21 serves as a molecular switch that controls the signal for cell growth and proliferation. The GTP-bound form of this enzyme turns the signal 'on' whereas the GDP-bound form turns the signal 'off'. Isolated ras p21 converts GTP to GDP in a rather slow intrinsic reaction ($k_{\text{cat}} = 0.028 \text{ min}^{-1}$; ref. 1). However, when Ras p21 interacts with the GTPase activating protein (GAP), catalysis is enhanced more than 10^5 -fold (ref. 2), turning off the growth signal. Mutation at either one of positions 12, 13 or 61 drastically reduces the catalytic effect of this enzyme, leaving the signal on and leading to uncontrolled cell growth.

The elucidation of the X-ray structure of Ras p21¹⁻⁷ has located all the residues which are involved in oncogenic mutations to the immediate vicinity of the γ -phosphate, but it seems that the correlation of structure and function of the isolated p21 protein is far from simple. The crystallographic studies identified a water molecule in a position that makes it likely to act as the nucleophile in the hydrolysis reaction. This water molecule is located precisely between Gln 61 and the γ -phosphate (Fig. 1). The steric arrangement of the active site led to the suggestion that Gln 61 plays an important part by accepting a proton from the nucleophilic water molecule⁷ (this mechanism is referred to here as the General Base 61 (GB61) mechanism). Recent theoretical⁸ and experimental^{9,10} studies have, however, indicated that the GB61 mechanism is unlikely to be the actual catalytic mechanism for the intrinsic hydrolysis reaction. It appears that the protein is unable to raise the pK_a of Gln 61 to a level that will catalyze the proton-transfer step. In fact, no other basic residue seems to be close enough to

act as a general base (Fig. 1). Recent crystal structures of GTPases related to Ras p21, — EF-Tu¹¹ and the α -subunit of transducin¹² — find a water molecule in a position very similar to that in p21 but with no interacting protein side chain. We have, therefore, to examine alternative mechanisms that do not involve a proton transfer to a general base, as well as mechanisms where the general base is the GTP itself or other nearby water molecules.

Alternative mechanisms

We used computer simulation to explore the validity of the most likely hydrolytic mechanisms of the intrinsic reaction of Ras p21. Our working hypothesis is that a mechanism can be eliminated if its activation barrier is drastically larger than the corresponding observed barrier, which was estimated to be around 23 kcal mol^{-1} (using the observed rate constant and the transition state theory⁸).

As in previous studies^{8,13,14}, we used the empirical valence bond (EVB) approach, a combined energy perturbation (FEP) and umbrella sampling approach as implemented in the program ENZYME¹⁵. This method describes a chemical reaction as a transition between different resonance structures. The potential surface for the given enzymatic reaction is calibrated by considering the corresponding mechanism in water as a reference reaction. This calibration procedure provides what is probably the most reliable current approach for evaluating the energetics of enzymatic reactions. Details of the simulation approach are given in the Methods section and in refs 13–15.

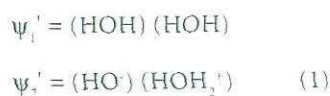
Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, USA

Correspondence should be addressed to A.W.

The results of the different simulation studies and the corresponding energy considerations are summarized below.

The role of a second water

X-ray studies of Ras p21^{7,9} located a second water molecule (Wat 189), near the nucleophilic water (Wat 175). Thus, it is possible that these two water molecules could undergo autoprotolysis and generate the activated OH⁻ nucleophile. Such a mechanism seems to be of importance in the catalytic reaction of carbonic anhydrase¹⁶, where proton transfer between two active site water molecules and a subsequent transfer of the proton from the active site to the surrounding solution has been found to be the rate-limiting step¹⁷. A proton transfer reaction between two water molecules can be described as an effective two-state problem corresponding to the valence bond (VB) structures: (1)



where in our specific case the two resonance structures (ψ_1 and ψ_2 in Fig. 2) also include the ROPO_3^{2-} fragment. The potential surface for the proton transfer step in the protein can be calibrated by considering the corresponding process in water as a reference. The energetics of the

reference reaction in water is known experimentally¹⁸ and is given by

$$\begin{aligned}\Delta G_{1 \rightarrow 2}^w &\sim 2.3 \text{ RT } [\text{pK}_a(\text{H}_2\text{O}) - \text{pK}_a(\text{H}_3\text{O}^+)] \sim 22 \text{ kcal mol}^{-1} \\ \Delta g_{1 \rightarrow 2}^* &\sim 25 \text{ kcal mol}^{-1}\end{aligned}\quad (2)$$

where w designates the reaction in water. In analysing the two water mechanism we have to consider a second step where the OH⁻ ion attacks the γ -phosphate. The reference reaction for this step is calibrated by the relationships:

$$\Delta G_{2 \rightarrow 3}^* = \Delta G[\text{OH}^- + \text{ROPO}_3^{2-} \rightarrow \text{ROP}(\text{O}_3^-)(\text{OH})] \sim 37 \pm 4 \text{ (kcal mol}^{-1}\text{)}$$

$$\Delta g_{2 \rightarrow 3}^* \sim \Delta G^*[\text{OH}^- + \text{ROPO}_3^{2-} \rightarrow \text{ROP}(\text{O}_3^-)(\text{OH})] \sim 46 \text{ kcal mol}^{-1}\quad (3)$$

where the observed values are taken from Guthrie's work¹⁹.

Fig. 2 presents the results of the simulation for the reference solution reaction and for the actual reaction in the protein active site. The calculated activation barrier is $\approx 68 \text{ kcal mol}^{-1}$ while the observed activation barrier for the intrinsic reaction of ras is around $\approx 23 \text{ kcal mol}^{-1}$ (see above). Thus, it is unlikely that the intrinsic reaction of ras involves a proton transfer between two

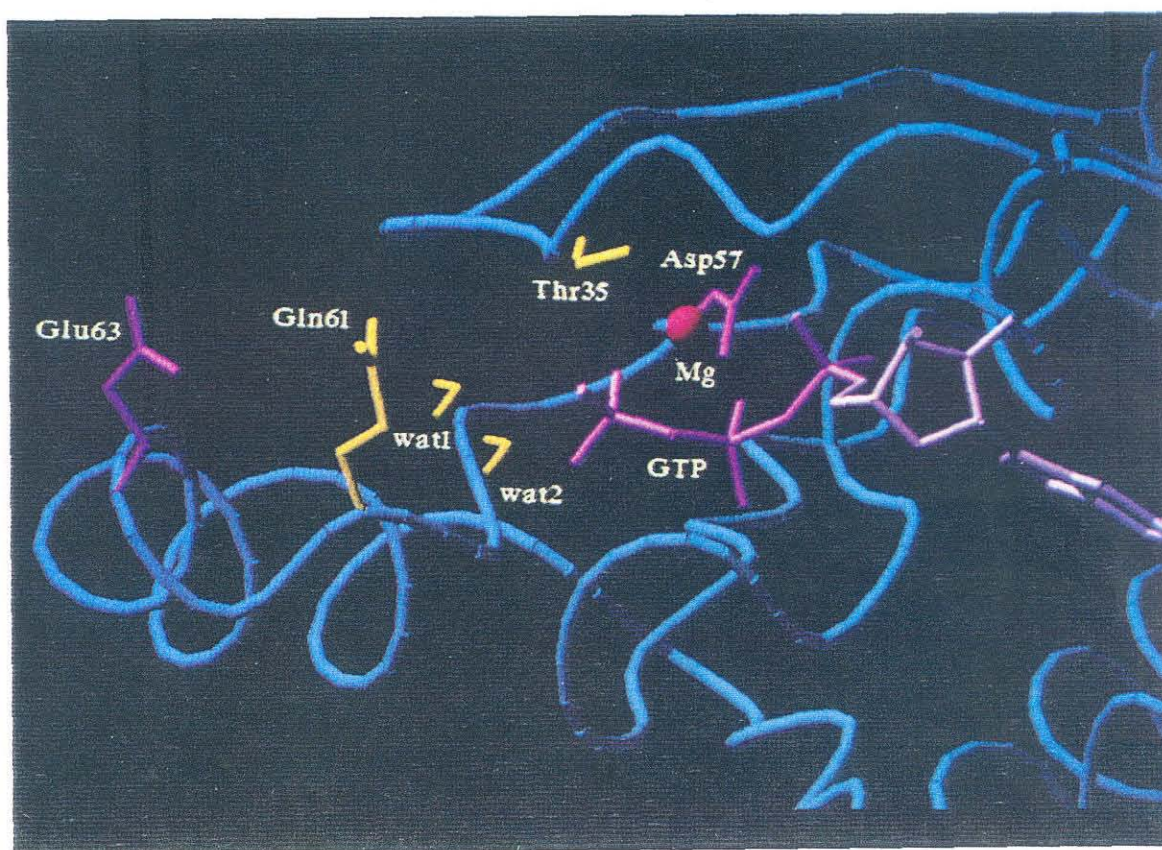


Fig.1 Good (purple) and bad (yellow) bases in the active site of ras p21.

water molecules. Note that the proton transfer step in the protein costs about ~ 49 kcal mol $^{-1}$ whereas the corresponding step in water costs only 22 kcal mol $^{-1}$. Apparently, the OH $^-$ H $_3$ O $^+$ ion pair is destabilized in the protein even more than the corresponding GlnH $^+$ OH $^-$ ion pair in the GB61 mechanism 8 . This destabilization is due in part to the repulsion between the OH $^-$ and the negatively charged GTP whose charges are not screened sufficiently by the protein environment. Furthermore, the H $_3$ O $^+$ ion is not stabilized in the protein as much as in water.

Neutral water attack

Considering the 'anticatalytic' effect of the electrostatic repulsion between the γ -phosphate and an attacking OH $^-$ ion, it is tempting to consider a mechanism where the nucleophile is a neutral water molecule. We considered such a mechanism by describing the system in terms of three resonance structures (Ψ_1, Ψ_2 and Ψ_3 of Fig. 3). The

potential surface of the reference reaction was calibrated using the following considerations: The free energy difference between Ψ_1 and Ψ_3 is given by 20 :

$$\Delta G_{1 \rightarrow 3}^w \sim 34 \text{ kcal mol}^{-1} \quad (4)$$

Furthermore, the free energy $\Delta G_{5 \rightarrow 4}^w$ can be estimated by considering the deprotonation of $\text{ROP}(\text{OH})_2(\text{O}_2^{2-})$ and the protonation of the (OH) group of $\text{ROP}(\text{OH})\text{O}_2^{3-}$. The free energy of the first step in this cycle is given by $\Delta G_s = 2.3 \text{ RT} \{(\text{pK}_s[\text{ROP}(\text{OH})_2(\text{O}_2^{2-}) - \text{pH}])\}$ (ref.21) whereas the free energy of the second step is given by $\Delta G_b = -2.3 \text{ RT} \{(\text{pK}_b[\text{ROP}(\text{OH})_2(\text{O}_2^{2-}) - \text{pH}])\}$. Thus we have

$$\begin{aligned} \Delta G_{5 \rightarrow 4}^w &= \Delta G_s + \Delta G_b \sim 2.3 \text{ RT} \{ \text{pK}_s[\text{ROP}(\text{OH})_2(\text{O}_2^{2-})] - \\ &\quad \text{pK}_b[\text{H}_3\text{O}^+] \} + [V_{\text{QO}}/\epsilon] \\ &\sim 26 - 12 = 14 \pm 10 \text{ kcal mol}^{-1} \quad (5) \end{aligned}$$

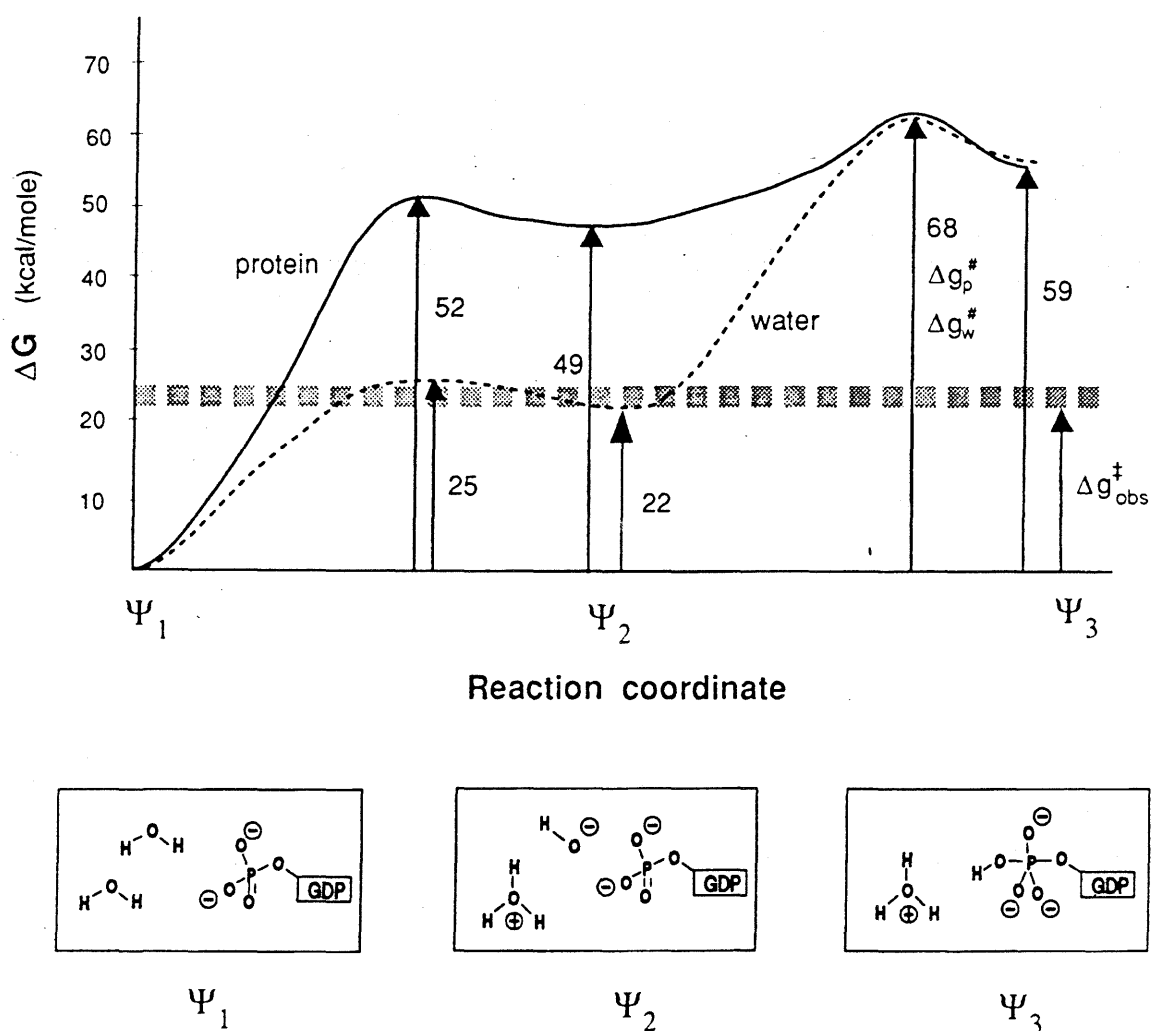


Fig.2 The EVB/FEP free-energy profiles for a reaction involving a second water molecule as a general base. The energetics of the reaction are compared in Ras p21 and in a reference solvent cage. The relevant resonance structures are shown below.

Here we use the fact that the pK_a of an H_2O^+ group which is bound to a neutral phosphate should be similar to the pK_a of H_3O^+ and estimate the electrostatic interaction between the H_2O^+ group and the negatively charged GTP by (V_{QQ}/ϵ) , where V_{QQ} is the gas phase interaction between the relevant charges and ϵ is the effective dielectric constant for this interaction. A dielectric constant of ~ 30 has been estimated by examination of zwitterion type model compounds²⁰ as well as from the analysis of the pK_a values of phosphorous acids.

Completing the thermodynamic cycle that involves the three resonance structures we obtain,

$$\Delta G_{1 \rightarrow 4}^* = \Delta G_{1 \rightarrow 5}^* + \Delta G_{5 \rightarrow 4}^* \sim 47 \text{ kcal mol}^{-1} \quad (6)$$

This estimate of $\Delta G_{1 \rightarrow 4}^*$ for the reference reaction is larger than 30 kcal mol^{-1} even when ϵ is reduced to the lowest limit of 15. As the activation barrier of the reaction in the protein is significantly higher than the corresponding barrier in solution, it is unlikely that such a reaction mechanism is involved in GTP hydrolysis in Ras p21.

Concerted mechanisms

One might argue that the reaction could proceed in a concerted fashion where two or more water molecules are involved in the hydrolysis. This would mean that bond breakage and bond formation occur at the same time. The feasibility of such concerted mechanisms was

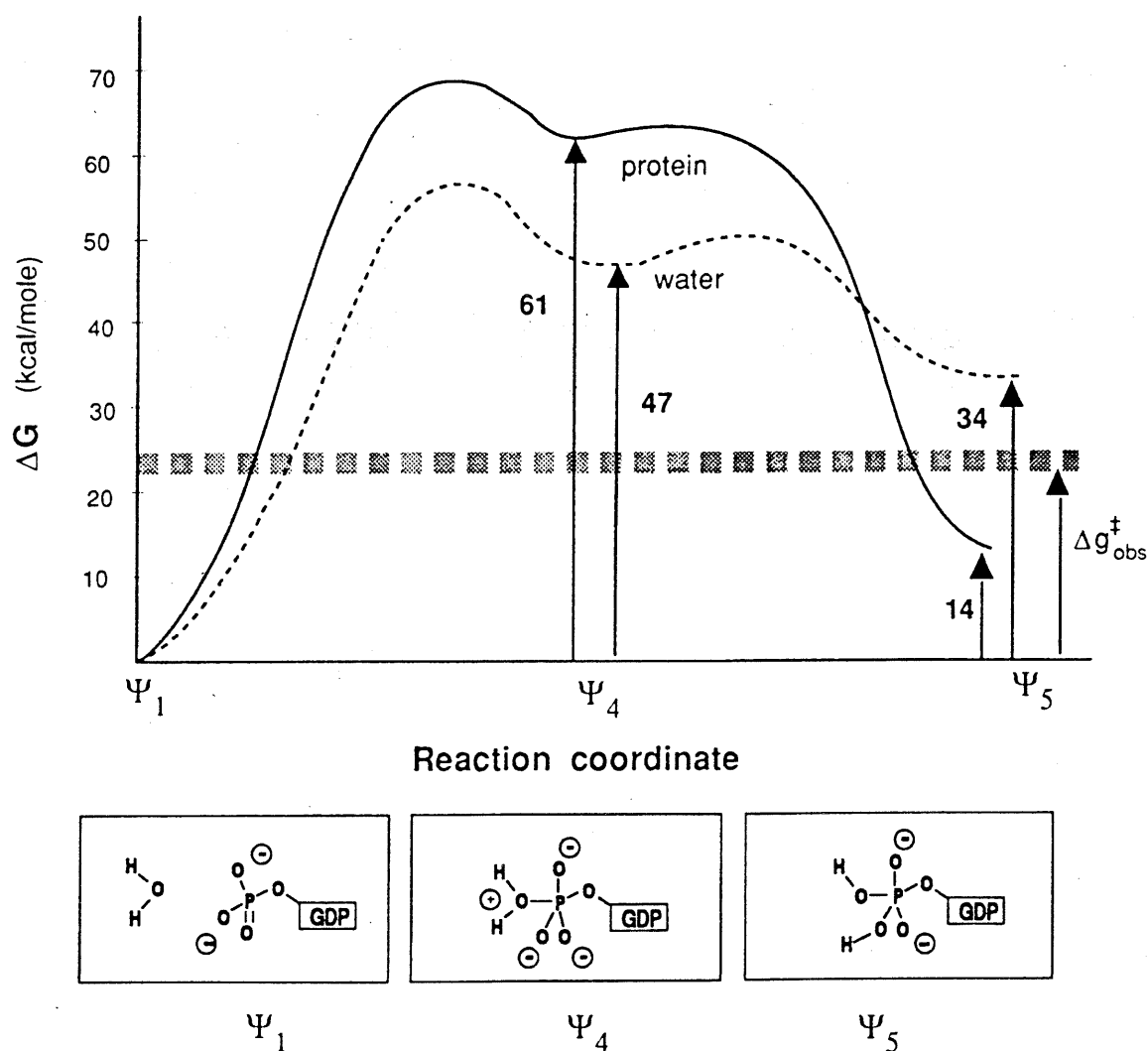


Fig.3 The free energy-profile for the attack of a single water molecule on the GTP γ -phosphate. Only the first step of this process is considered, as the corresponding barrier is already much too high to account for the observed rate constant of the intrinsic reaction of Ras p21. The FEP/EVB calculations for the reaction in the protein and in a reference solvent cage are compared. The relevant resonance structures are shown below.

examined by evaluating the corresponding contour maps using the resonance structures Ψ_1, Ψ_2, Ψ_3 and Ψ_4 and alternatively Ψ_1, Ψ_2, Ψ_3 and Ψ_5 . The reference reactions were calibrated by the same information used in the simulations of the stepwise pathways that defined the sides of the concerted surfaces (Figs 4,5). In the case of Fig. 5, we also considered in the calibration procedure the path from Ψ_2 to Ψ_3 using: (7)

$$\Delta G_{2 \rightarrow 3} = \Delta G_{2 \rightarrow 4} - 2.3RT \{pK_a[\text{ROP}(\text{OH})_2(\text{O}_2^{2-}) - pK_a[\text{H}_2\text{O}^+]] \sim 11 \text{ kcal mol}^{-1}$$

The mapping procedure used to obtain free energy surface is discussed in the method section.

The activation barriers for the concerted paths are

higher than 40 kcal mol^{-1} , a value significantly larger than the observed barrier for ras p21 (Figs 4,5). Thus, it seems unlikely that the intrinsic reaction of ras involves a concerted attack involving two water molecules.

γ -phosphate acts as a general base

In view of the above results it is hard to rationalize the catalytic effect of ras p21. Yet there is one general base that we did not consider, namely the γ -phosphate. The feasibility of a mechanism in which the GTP is the general base for its own hydrolysis (which is referred to here as the GBy mechanism), was examined by calibrating the potential surface of the corresponding reference reaction, using the following relationships

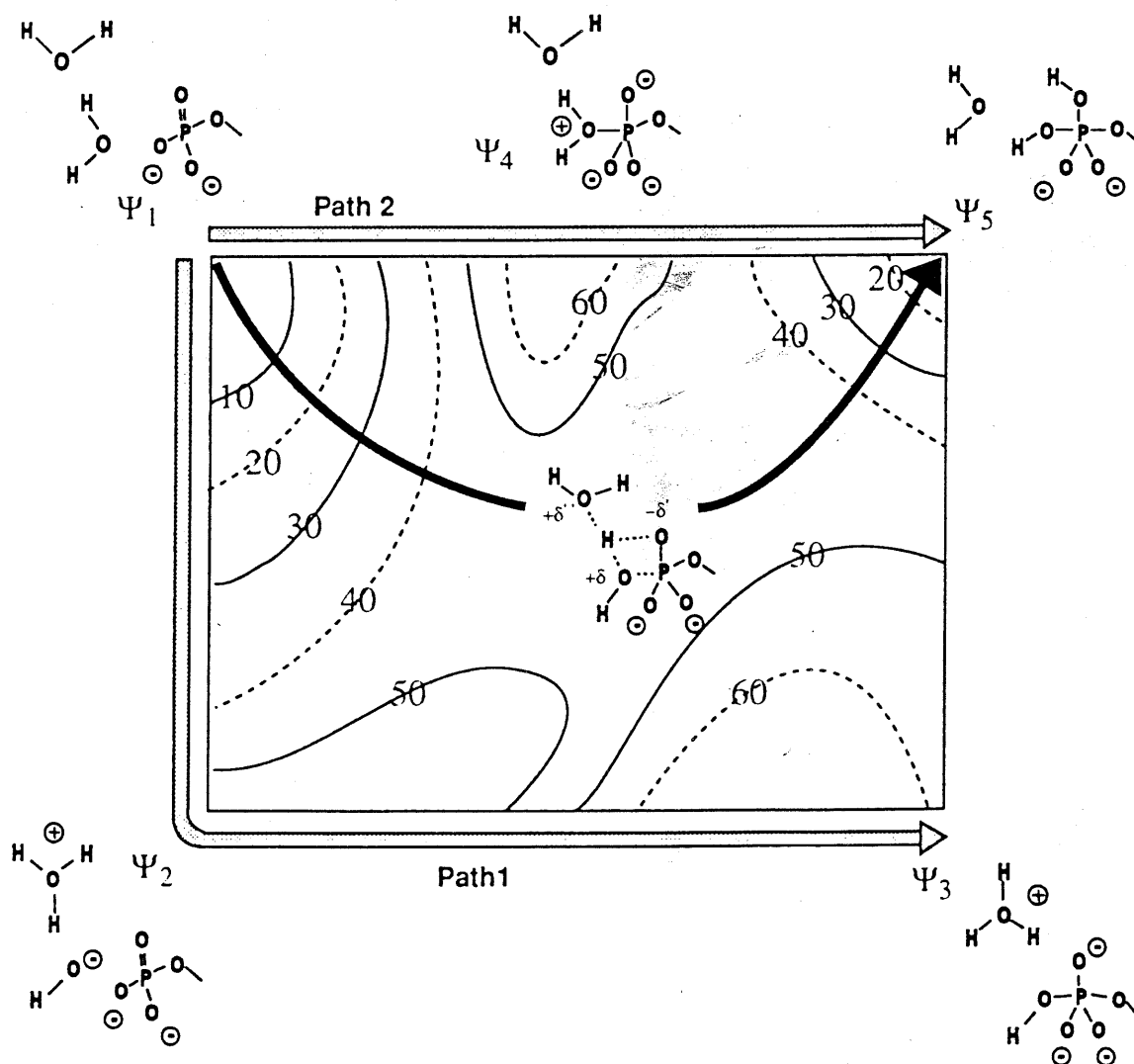


Fig.4 The free-energy surface for a concerted mechanism that involves Ψ_1, Ψ_2, Ψ_3 and Ψ_4 . Note that the stepwise mechanisms of paths 1 and 2 correspond to the reactions presented in Figs 2 and 3, respectively.

$$\Delta G_{1 \rightarrow 6}^* \sim -2.3 RT [\text{p}K_a(\text{ROP}(\text{O})_2\text{H}(\text{OH})) - \text{p}K_a[\text{H}_2\text{O}]] \sim 12 \text{ kcal mol}^{-1}$$

$$\Delta G_{6 \rightarrow 5}^* = \Delta G[\text{OH}^- + \text{ROP}(\text{O})_2\text{H}(\text{OH}) \rightarrow \text{ROP}(\text{O})_2\text{H}(\text{OH})_2] \sim 22 \text{ kcal mol}^{-1}$$

$$\Delta G_{6 \rightarrow 4}^* = \Delta G[\text{H}^+ + \text{ROP}(\text{O})_2\text{H}(\text{OH}) \rightarrow \text{ROP}(\text{O})_2\text{H}(\text{OH})_2] \sim 34 \text{ kcal mol}^{-1} \quad (8)$$

where the value of $\Delta G_{6 \rightarrow 5}^*$ was taken from ref.19 and $\Delta G_{6 \rightarrow 4}^*$ was estimated from the corresponding rate constant.

The calculated free-energy surfaces for the GBy mechanisms in the protein and in solution are presented in Fig. 6, which shows that the activation energy in the protein is much lower than it is in the other mechanism.

Furthermore, the overall activation free energy for both steps is $24 \pm 5 \text{ kcal mol}^{-1}$, in good agreement with the observed activation free energy of ras p21. The reasons for the reduction in the overall barrier are quite instructive. The proton transfer step costs less than in the other mechanisms both because of the higher $\text{p}K_a$ of the base and the fact that the protonation of the γ -phosphate reduces the destabilization of the OH^- ion. The nucleophilic attack step also involves a lower barrier as the protonation of the γ -phosphate reduces the barrier for this process and as the negative charge that comes to reside on the γ -phosphate is stabilized by the electrostatic po-

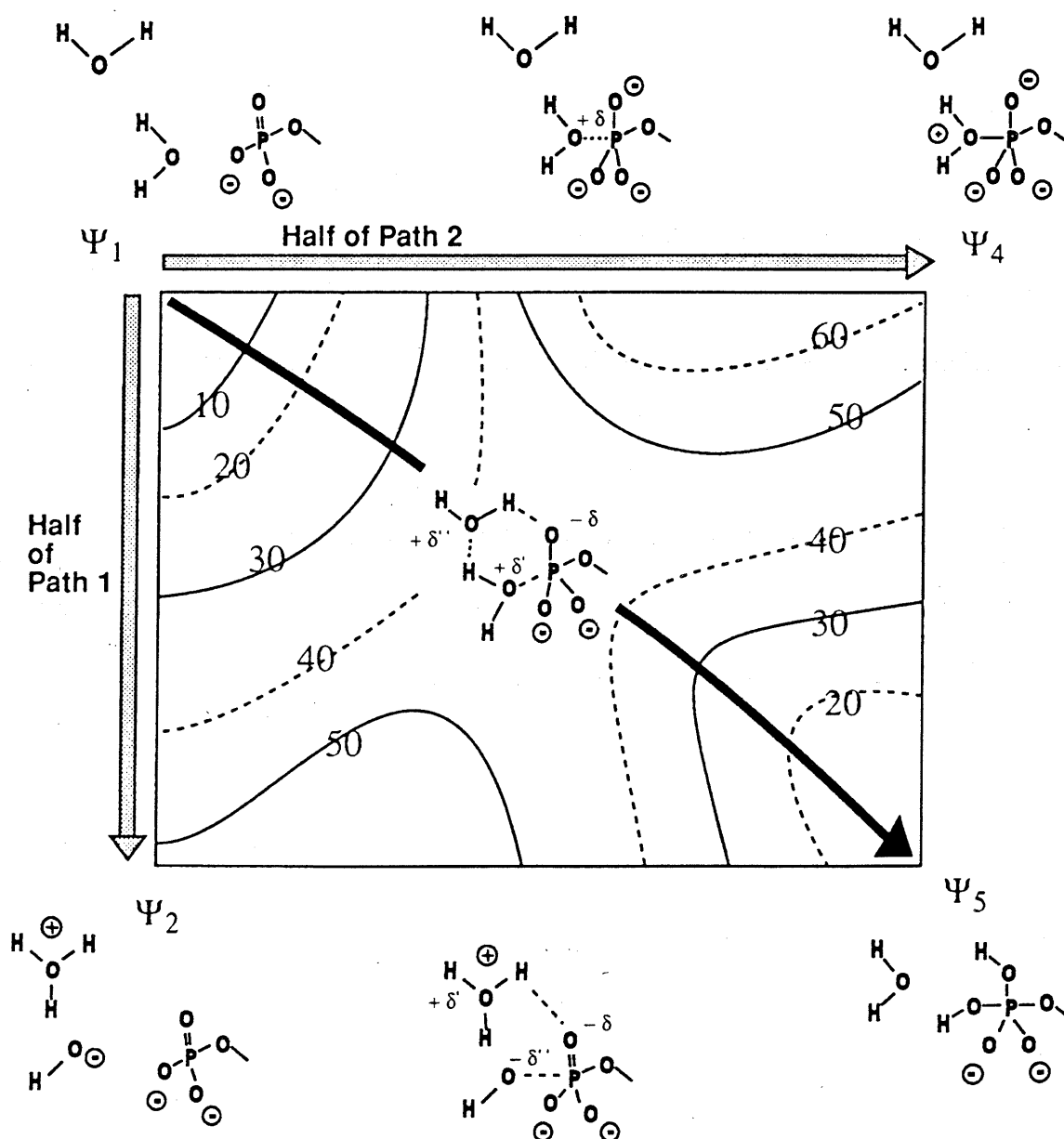


Fig. 5 The free-energy surface for a concerted mechanism that involves Ψ_1 , Ψ_2 , Ψ_5 and Ψ_4 .

tential of the Mg^{2+} ion and Lys 16. This catalytic effect is also involved in the second step of the mechanisms depicted in Figs 2 and 3 (for example, the free energy difference between ψ_2 and ψ_3 in Fig. 2) but it does not help to reduce the very high barrier in the initial step of each of these mechanisms. We conclude that the GBy mechanism is likely to be the actual mechanism of the intrinsic reaction of Ras p21.

Other options

This study has not considered a mechanism that involves Asp 57 as a general base for a water molecule that is bound to the Mg^{2+} ion, although such a mechanism may be operative in staphylococcal nuclease at low pH¹⁴. However in ras p21 the nucleophilic attack of such a water molecule is inconsistent with the observation of in-line inversion of the γ -phosphate centre²². Furthermore, site-directed mutagenesis experiments in EF-Tu have shown that the homologous mutation Asp to Asn leads to an increased hydrolysis rate²³.

Another related mechanism that is not simulated in the present study, is a specific base catalysis mechanism where an OH^- ion from the solution around the protein serve as the base in the reaction. In this case, the free energy for the formation of the OH^- ion from the catalytic water (or for a replacement of this water by an OH^- ion) is given by²¹.

$$-\Delta G = 2.3 RT [pK_a^f(H_2O) - pH] \quad (9)$$

This mechanism should result in an increase of the rate constant with increased pH of the solution around the protein, as it is the case in the reaction of the exo site of DNA polymerase (A.W., *et al.*, manuscript in preparation). But such a relationship has not been observed experimentally²⁴. Furthermore, the reduction of the pK_a of the nucleophilic water molecule usually requires a nearby metal ion and this excludes all the active site water molecules except those bound to the Mg^{2+} ion. However, as stated above the nucleophilic attack of this water is inconsistent with experimental observations.

The reaction of ras p21 might involve a dissociative mechanism where the bond to the γ -phosphate breaks before or in concert with the nucleophilic attack step (see for example ref. 25). However, a fully dissociative mechanism is unlikely, as the Mg^{2+} ion is probably designed to stabilize the migration of a negative charge to the GTP. Moreover, a concerted associative–dissociative mechanism will be correlated linearly with the energetics of the nucleophilic attack pathway that was studied in this work.

The GBy mechanism

The GBy mechanism can explain why mutagenesis studies involving different active-site residues did not result in a significant change in the intrinsic rate of ras p21. The GBy mechanism has also several advantages over the other mechanisms considered here. The γ -phosphate has the highest pK_a of all the bases analyzed. The protonation of the γ -phosphate considerably reduces the barrier for the nucleophilic attack as the OH^- ion encounters less electrostatic repulsion during this step. This mechanism

avoids the formation of high-energy intermediates of the type shown in Figs 2, 3, 4 and 5.

The proposed mechanism can be verified by simulating the effect of mutations on the reaction rate, and such a study is now in progress. Another aspect of our calculation should also be testable. We predict that the free energy for the proton transfer to the γ -phosphate in the protein is larger than in water. Analysis of our calculations has revealed that this increase is mainly caused by a drop in the pK_a of the γ -phosphate to around three units (due in part to the effect of the Mg^{2+} ion), whereas the pK_a change for the catalytic water molecule is less pronounced. The prediction of a downshift in the pK_a of the γ -phosphate can be tested experimentally.

If the GBy mechanism is operative in Ras, then 'mutations' at the γ -phosphate itself would be expected to lead to a significant change in the hydrolysis rate. Of course, the analysis of the results of such experiments would not be easy as they would involve a simultaneous change of the pK_a of the proton acceptor, the strength of the leaving group and the interaction with the protein. However, computer simulation studies should allow one to evaluate the difference between the rate of the reaction of a given substrate in p21 and in water and to determine whether the observed effect is consistent with the GBy mechanism.

We have examined here the intrinsic mechanism of ras p21 and not the action of the ras/GAP complex. Nevertheless it is possible that the role of GAP is to catalyze the GBy-reaction rather than to promote another reaction mechanism. In this case it might be possible that GAP helps in optimizing the interaction between Gln 61 and the OH^- nucleophile. It is also possible that GAP provides some residues that stabilize the pentacoordinated transition state.

Methods

In order to examine different mechanisms by a theoretical approach, the approach must be sufficiently accurate to eliminate incorrect mechanisms. At present, one of the most reliable methods is the EVB approach, which considers the energetics of enzymatic reactions relative to closely related processes in solutions. The method describes the given reaction in terms of resonance structures that represent the bonding and charge distribution of different feasible valence bond states. The energies of the resonance structures are represented by:

$$E = U_{bond}^{(1)} + U_{bending}^{(2)} + U_{torsion}^{(3)} + U_{stretch}^{(4)} + U_{solv}^{(5)} + U_{other}^{(6)} \quad (10)$$

where the first four terms describe the energies of the molecular fragments that correspond to the specific bonding and charge distribution of the i^{th} resonance structure, the fifth term represents the interaction between the reaction system (the solute) and the surrounding solvent, while the sixth term represents the solvent–solute interaction. The parameter α is the so-called 'gas phase shift' that determines the equilibrium energy of the i^{th} form in the gas phase. The actual ground state energy of the system, E_g , is determined by mixing the VB states using the relevant off-diagonal matrix elements (the H_{ij} elements). The details of the EVB approach, including the functional forms of the ϵ_i values and the H_{ij} values are given elsewhere^{15,27}. The key point in the EVB procedure is the reliable calibration of the ϵ_i and H_{ij} values using experimental information about the reaction free energy and activation energies in solution (see ref. 13 for details). After calibrating the potential surfaces for the reference solution reactions, we use the relevant parameters (e.g. α) without any change for studying the

corresponding reaction in the protein, where the only change is the substitution of the solute-solvent environment by the solute-protein surrounding in Eq. (10) (U_{ss} becomes U_{ps}). The free-energy surface defined by the EVB ground state energy E_g is evaluated by a combination of a free-energy perturbation and umbrella sampling method described elsewhere¹³.

The parameters used in the present work for Eq. (10) are summarized in ref. 8. The simulations were performed with the ENZYME program¹⁵ using a surface-constrained model¹³. The fragments shown explicitly in Figs 2-6 were represented by the VB force field and the rest of the system was represented by a standard force field.

The protein atoms at a distance of 20 Å or more from the γ-phosphate were constrained to stay near the X-ray structure provided by S.H. Kim. The active site was solvated by a 15 Å sphere of explicit water molecules (which involves a surface region that is subject to a special set of constraints¹⁵) and this water sphere was further extended to 25 Å by surrounding it with Langevin dipoles. The free-energy calculations were done using a mapping potential of the form

$$\epsilon_m = (1 - \lambda_m)\epsilon_a + \lambda_m\epsilon_c \quad (11)$$

where ϵ_a and ϵ_c are the energies of the initial and final resonance structures of the given reaction step (e.g. ψ_1 and ψ_3 in Fig. 2). The simulations involve a 15 ps equilibration followed by a change of λ_m from zero to one in ten steps of 4 ps each. The free energy of the ground state was evaluated in terms of a reaction that corresponds to the energy difference between different resonance structures as described elsewhere¹⁵.

The evaluation of the free-energy surfaces for the concerted pathways involved the following procedure. The surface was defined in terms of four resonance structures a, b, c and d, where a and d defined the upper panel of the surface (e.g. ψ_1 and ψ_3 of Fig. 4) while b and c defined the lower panel (e.g. ψ_2 and ψ_4 in Fig. 4). The free energy surface was then mapped with a potential of the form

$$\epsilon_{m,n} = [(1 - \lambda_m)\epsilon_a + \lambda_m\epsilon_b](1 - \lambda_n) + [(1 - \lambda_m)\epsilon_c + \lambda_m\epsilon_d]\lambda_n \quad (12)$$

The mapping procedure was done by changing λ_n in five steps and the reaction coordinate was defined in terms of $(\epsilon_b - \epsilon_a)$ and $(\epsilon_d - \epsilon_c)$.

Received 22 December 1993, accepted 6 May, 1994

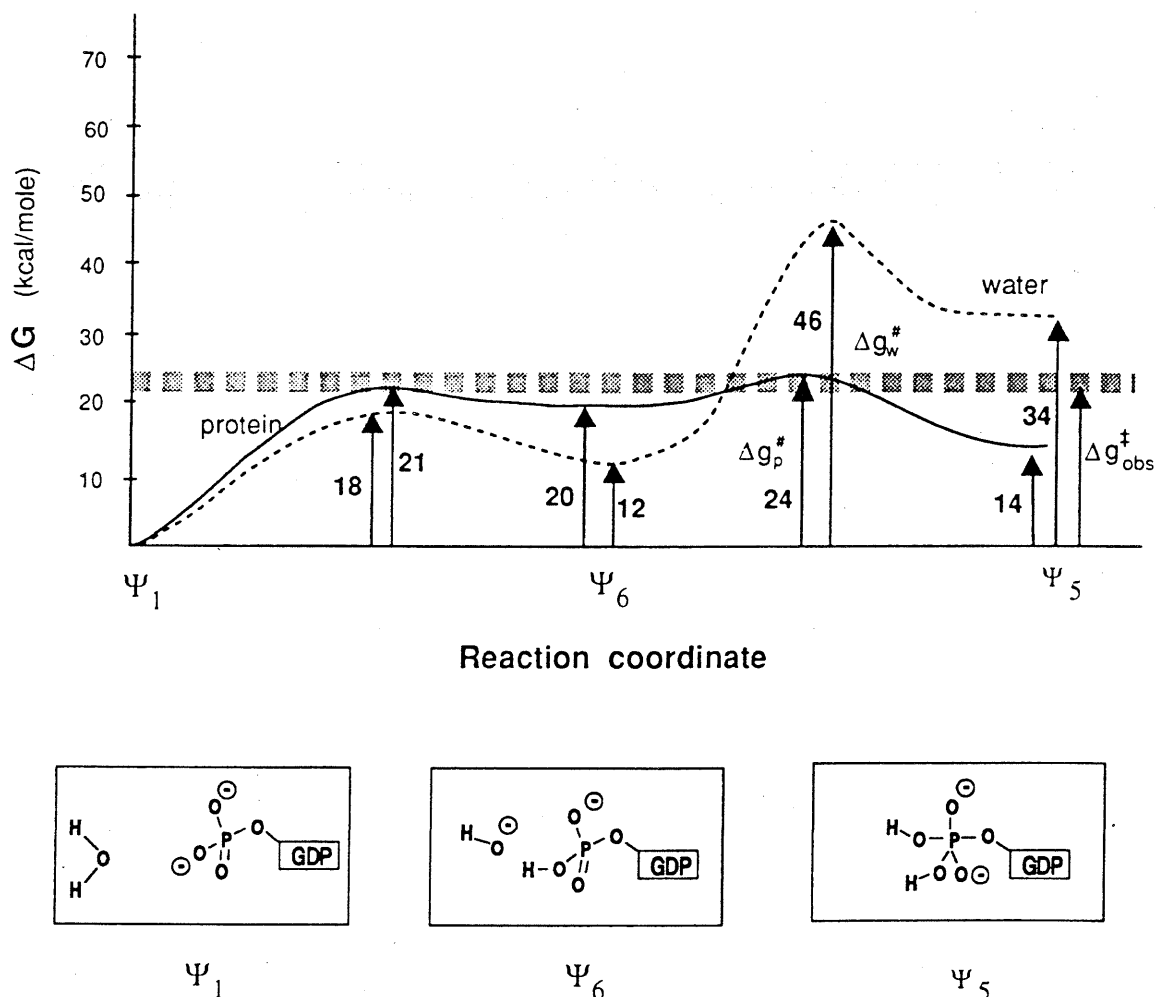


Fig. 6 EVB/FEP free-energy profiles for a reaction involving the γ-phosphate of the GTP as a general base. The profile for the reaction in ras p21 is compared with that in the reference solution reaction.

article

Acknowledgements

We thank S.H. Kim for providing the coordinates of p21. This work was supported by a grant from the NIH to A.W. T.S. was supported by from the Gottlieb Daimler and Karl Benz Foundation. R.L. was supported by the Charles Heidelberger Memorial Foundation.

1. Temeles, G.L., Gibbs, J.B., D'Alonzo, J.S., Sigal, I.S. & Scolnick, E.M. Yeast and mammalian Ras proteins have conserved biochemical properties. *Nature* **313**, 700–703 (1985).
2. Gideou, P. et al. Mutational and kinetic analysis of the GTPase-activating protein (GAP)-p21 interaction. *Molec. Cell. Biol.*, **12**, 2050–2054 (1992).
3. De Vos, A.M. et al. 3-Dimensional structure of an oncogene protein - catalytic domain of human c-H-Ras. *Science* **239**, 888–893 (1988).
4. Krengel, U. et al. 3-Dimensional structures of H-Ras p21 mutants - molecular basis for their inability to function as signal transduction switch molecules. *Cell* **62**, 539–548 (1990).
5. Milburn, M. et al. Molecular switch for signal transduction - structural differences between active and inactive forms of protooncogenic Ras mutants. *Science* **247**, 939–945 (1990).
6. Pai, E.F. et al. Structure of the guanine-nucleotide-binding domain of the Ha-Ras oncogene product p21 in the triphosphate conformation. *Nature* **341**, 209–214 (1989).
7. Pai, E.F. et al. Refined crystal-structure of the triphosphate conformation of H-Ras p21 at 1.35Å resolution—implications for the mechanism of GTP hydrolysis. *EMBO J.* **9**, 2351–2359 (1990).
8. Langen, R., Schweins, T. & Warshel, A. On the mechanism of guanosine triphosphate hydrolysis in Ras p21 proteins. *Biochemistry* **31**, 8691–8696 (1992).
9. Prive, G.G. et al. X-Ray crystal-structures of transforming p21 Ras mutants suggests a transition-state stabilization mechanism for GTP hydrolysis. *Proc. natn. Acad. Sci. U.S.A.* **89**, 3649–3653 (1992).
10. Chung, H.-H., Benson, D.R., & Schultz, P.G. Probing the structure and mechanism of Ras protein with an expanded genetic code. *Science* **259**, 806–809 (1993).
11. Berchtold, H. et al. Crystal-structure of active elongation-factor Tu reveals major domain rearrangements. *Nature* **365**, 368–368 (1993).
12. Noel, J.P., Hamm, H. & Sigler, P.B. The 2.2Å crystal structure of transducin complexed with GTPγS. *Nature* **366**, 654–663 (1993).
13. Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (John Wiley, New York, 1991).
14. Aqvist, J. & Warshel, A. Calculations of the free energy profiles for the staphylococcal nuclease catalyzed reactions. *Biochemistry* **28**, 4680–4689 (1989).
15. Lee, F.S., Chu, Z.T. & Warshel, A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIK programs. *J. comp. Chem.* **14**, 161–185 (1993).
16. Aqvist, J. & Warshel, A. Computer simulation of the initial proton-transfer step in human carbonic anhydrase. *J. molec. Biol.* **224**, 7–14 (1992).
17. Silverman, D.N. & Lindskog, S. The catalytic mechanism of carbonic anhydrase- Implications of a rate-limiting protolysis of water. *Accts chem. Res.* **21**, 30 (1988).
18. Eigen, M. & de Mayer, L. Kinetics of neutralization Z. *Elektrochem.* **59**, 986–993 (1955).
19. Guthrie, J.P. Hydration and dehydration of phosphoric acid derivatives: free energies of formation of the pentacoordinate intermediates for phosphate ester hydrolysis and of monomeric metaphosphate. *J. Am. chem. Soc.* **99**, 3991–4001 (1977).
20. Warshel, A. & Russell, S. Calculations of electrostatic interactions in biological systems and in solution. *Q. Rev. Biophys.* **17**, 283–422 (1984).
21. Warshel, A. Calculations of enzymatic reactions: calculations of pKa, proton transfer reactions, and general acid catalysis reactions in enzymes. *Biochemistry* **20**, 3167 (1981).
22. Feuerstein, J., Goody, R.S. & Webb, M.R. The mechanism of guanosine nucleotide hydrolysis by p21 c-Ha-Ras. *J. biol. Chem.* **264**, 6188–6190 (1989).
23. Harmark, K., Anborgh, P.H., Merola, M., Clark-B.F.C. & Parmeggiani, A. Substitution of aspartic acid-80, a residue involved in coordination of magnesium, weakens the GTP binding and strongly enhances the GTPase of the G-domain of elongation factor-Tu. *Biochemistry* **31**, 7367–7372 (1992).
24. Mistou, M.Y., Cool, R.H. & Parmeggiani, A. Effects of ions on the intrinsic activities of C-H-Ras protein p21- a comparison with elongation factor-Tu. *Eur. J. Biochem.* **204**, 179 (1992).
25. Williams, A. Concerted mechanisms of acyl group transfer reactions in solution. *Accts. chem. Res.* **22**, 387–392 (1989).
26. Warshel, A., Sussman, F., & Hwang, J.-K. Evaluation of catalytic free energies in genetically modified proteins. *J. molec. Biol.* **201**, 139–159 (1988).